Collection M / Sciences fondamentales

Analyse Hilbertienne

Houcine Chebli



Centre de Publication Universitaire

ANALYSE HILBERTIENNE

Par

Houcine Chebli

Professeur de Mathématiques

Faculté des sciences de Tunis

Centre de Publication Universitaire

Tunis, 2001

AVANT PROPOS

Le contenu de cet ouvrage correspond ` a l'enseignement dispensé aux étudiants de deuxième année de ma^ıtrise de mathématiques.

Le livre s'adresse principalement aux étudiants et à ce titre, je n'ai pas hésité à détailler les démonstrations des résultats énoncés, même lorsque parfois elles semblent évidentes. Cela m'a amené à me répéter de temps à autre. Les théorèmes utilisés sans démonstration, sont rappelés en annexe avec leurs démonstrations. Cela facilitera la lecture de l'ensemble sans autre recours, mais ne doit pas priver l'étudiant de lire d'excellents ouvrages sur le sujet, comme par exemple celui de A. Achour " Calcul Différentiel" paru dans la même collection et où on trouve des exercices intéressants.

Comme la plupart des résultats présentés ici sont classiques, je n'ai pas essayé de me renseigner sur la source de chaque partie. J'ai simplement signalé dans la bibliographie les principaux livres que j'ai consultés. En aucun cas l'absence de référence n'implique une prétention à l'originalité.

Le livre est organisé en cinq chapitres et une annexe. Chaque chapitre est divisé en sections et chaque section est suivie d'une liste d'exercices qui sont souvent une application directe du cours. On a proposé des solutions à certains d'entre eux; ces solutions ne sont pas toujours détaillées et ne constituent pas un modèle du genre, l'objectif visé est d'encourager le lecteur à traiter tous les exercices ce qui lui permettra de contrôler la bonne assimilation du cours.

Le texte a été réalisé en utilisant TEX. Cet incomparable traitement de texte est une arme à double tranchant : il permet de tout faire, mais dévoile facilement le mauvais goût. J'ai essayé, autant que faire se peut, de minimiser ces fautes de goût. Il n'en reste pas moins que utiliser TEX ne suffit pas à faire un bon et beau livre. J'espère ne pas trop décevoir ceux qui voudrons bien lire ce livre.

Le referee m'a fait part de beaucoup de corrections, critiques et commentaires, c'est un plaisir pour moi de lui exprimer mes remerciements sincères pour cette généreuse collaboration...

Février, 2001 — Houcine Chébli

Table des matières

Chapitre 1

Espace de Hilbert

Les espaces de Hilbert[1] sont la version de dimension infinie des espaces euclidiens ou hermitiens, dont ils gardent beaucoup de propriétés. En fait, ils trouvent leur origine dans la théorie des développement de fonctions arbitraires en séries de fonctions orthogonales, celles-ci apparaissant le plus souvent comme fonctions propres de certains opérateurs différentiels linéaires (séries de Fourier, fonctions sphériques, polynômes orthogonaux). Ils fournissent le cadre mathématique dans lequel se développe la mécanique quantique et jouent un rôle important dans beaucoup de branches des mathéma-tiques, spécialement en Analyse linéaire.

1.1 Propriétés élémentaires et Exemples.

Dans tout ce chapitre $\mathbb{K}$ désigne ou bien le corps des nombres réels $\mathbb{R}$ ou bien le corps des nombres complexes $\mathbb{C}$.

Définition 1.1.1. *Soit E un espace vectoriel sur $\mathbb{K}$. Une forme sesquilinéaire sur E, est toute application B de $E \times E$ dans $\mathbb{K}$ vérifiant, quels que soient α, β dans $\mathbb{K}$ et x, y, z dans E*

(a) $B(\alpha x + \beta y, z) = \alpha B(x, z) + \beta B(y, z)$

(b) $B(x, \alpha y + \beta z) = \overline{\alpha} B(x, y) + \overline{\beta} B(x, z)$

On dit que B est hermitienne si elle vérifie de plus

(c) $B(x, y) = \overline{B(y, x)}, \quad \forall\, x, y \in E$

Notons que dans le cas où le corps des scalaires est $\mathbb{K} = \mathbb{R}$, une forme sesquilinéaire est simplement une forme bilinéaire, et une forme hermitienne

[1]Le mathématicien allemand David HILBERT (1862-1943) est l'un des plus grands mathématiciens de son temps. Il a contribué à presque toutes les branches des mathématiques, de la logique à l'algèbre en passant par l'analyse et la géométrie. Lors du Congrès International des mathématiciens tenu à Paris en 1900, il a formulé 23 problèmes qui ont servi de référence dans la recherche mathématique et ouvert la voie à plusieurs générations de chercheurs.

est une forme bilinéaire symétrique. La forme quadratique associée à B est définie par : $Q(x) = B(x,x)$. On notera que si B est hermitienne alors Q est réelle, c'est-à-dire que $Q(x)$ est un réel pour tout x dans E.
On dit que la forme hermitienne B est positive si la forme quadratique Q est positive : $Q(x) \geq 0$, pour tout x dans E.

Théorème 1.1.2. (Identité de polarisation)
Toute forme bilinéaire symétrique B vérifie

$$4B(x,y) = B(x+y,x+y) - B(x-y,x-y)$$

Toute forme sesquilinéaire B (hermitienne ou non), vérifie

$$4B(x,y) = B(x+y,x+y) - B(x-y,x-y) + \\ + iB(x+iy,x+iy) - iB(x-iy,x-iy)$$

Ces identités, dont la preuve est immédiate, montrent que, dans tous les cas, la forme quadratique Q associée à B caractérise entièrement celle-ci.

Théorème 1.1.3. *Soit B une forme hermitienne sur E. Si B est positive alors elle vérifie les propriétés fondamentales suivantes*
(i) *L'inégalité de* Cauchy[2]-Schwarz[3]

$$|B(x,y)| \leq B(x,x)^{\frac{1}{2}} B(y,y)^{\frac{1}{2}}$$

L'égalité n'a lieu que si x et y sont proportionnels.
(ii) *L'inégalité de* Minkowski[4]

$$B(x+y,x+y)^{\frac{1}{2}} \leq B(x,x)^{\frac{1}{2}} + B(y,y)^{\frac{1}{2}}$$

Démonstration. Le point-clef est bien sûr la positivité de la forme hermitienne. Etant donné deux éléments x et y, on considère la fonction ϕ définie sur $\mathbb{K}$ par $\phi(\alpha) = B(x+\alpha y, x+\alpha y)$. C'est une fonction à valeurs positives et par développement, elle sécrit

$$\phi(\alpha) = |\alpha|^2 B(y,y) + \overline{\alpha} B(x,y) + \alpha B(y,x) + B(x,x)$$

[2]Augustin-Louis CAUCHY (1789-1857) compte d'importants travaux sur les intégrales définies, la théorie des fonctions d'une variable complexe et la théorie de l'élasticité. Il a donné une nouvelle architecture à l'Analyse.

[3]Karl Hermann Amandus SCHWARZ (1843-1921), mathématicien allemand d'une grande intuition géométrique, a notamment établi le théorème d'uniformisation pour les domaines simplement connexes du plan complexe.

[4]Hermann MINKOWSKI (1864-1921), est un mathématicien russo-allemand à l'esprit original et perspicace dont Einstein avait été l'étudiant au Polytechnicum de Zürich. Il a introduit l'idée fondamentale qui consiste à envisager l'espace et le temps comme formant une seule et même entité : *l'espace-temps à quatre dimensions.*

Supposons $B(y,x) = |B(x,y)|e^{i\theta}$ et soit $\alpha = te^{-i\theta}$, avec t dans $\mathbb{R}$, et soit $\phi(\alpha) = P(t)$. Ce qui précède montre que, pour tout réel t,

$$P(t) = t^2Q(y) - 2t|B(x,y)| + Q(x) \geq 0$$

Si $Q(y) = 0$ l'inégalité précédente ne peut avoir lieu pour tout réel t que si $B(x,y) = 0$ et l'inégalité de Cauchy-Schwarz est dans ce cas une égalité. Sinon, P est un polynôme du second degré qui reste positif en tout t réel ; il en résulte que son discriminant, à savoir $|B(x,y)|^2 - Q(y)Q(x)$, est négatif ce qui traduit précisément l'inégalité de Cauchy-Schwarz. De celle-ci on déduit

$$B(x,y) + B(y,x) = 2\Re e B(x,y) \leq 2B(x,x)^{\frac{1}{2}}B(y,y)^{\frac{1}{2}}$$

et en ajoutant $B(x,x) + B(y,y)$ aux deux membres, on trouve

$$B(x+y, x+y) \leq \left[B(x,x)^{\frac{1}{2}} + B(y,y)^{\frac{1}{2}}\right]^2$$

En prenant la racine carrée des deux membres, on obtient l'inégalité de Minkowski. □

L'inégalité de Minkowski montre que l'application $x \mapsto Q(x)^{\frac{1}{2}}$ est une semi-norme, et est une norme si et seulement si la forme Q est non dégénérée, c'est-à-dire vérifie : $Q(x) = 0$ si et seulement si $x = 0$. On dit, dans ce cas, que c'est une norme induite par la forme hermitienne B.

Définition 1.1.4. *Un produit scalaire sur E est une forme hermitienne positive et non dégénérée.*

Définition 1.1.5. *Un espace préhilbertien réel (ou complexe) est un espace vectoriel E sur $\mathbb{K}$ sur lequel est défini un produit scalaire B et muni de la norme induite par B.*

On écrit $\langle x,y\rangle$ au lieu de $B(x,y)$ et on pose $\|x\| = Q(x)^{\frac{1}{2}}$. Avec ces notations, les inégalités de Cauchy-Schwarz et de Minkowski s'écrivent, pour x et y dans E,

$$|\langle x,y\rangle| \leq \|x\|\,\|y\| \quad \text{et} \quad \|x+y\| \leq \|x\| + \|y\|$$

Exemple 1.1.6. - Les espaces $\mathbb{R}^n$ et $\mathbb{C}^n$ sont des espaces préhilbertiens pour le produit scalaire dit *usuel*

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{n} x_i y_i, \quad \text{dans le cas réel}$$

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{n} x_i \overline{y_i}, \quad \text{dans le cas complexe}$$

Exemple 1.1.7. - Soit $\ell^2(\mathbb{N})$ l'ensemble des suites $\mathbf{x} = (x_n)$, avec n dans $\mathbb{N}$ et x_n dans $\mathbb{K}$, telles que

$$\sum_{i=1}^{\infty} |x_i|^2 < \infty$$

Soient $\mathbf{x} = (x_n)$ et $\mathbf{y} = (y_n)$ deux éléments de l'ensemble $\ell^2(\mathbb{N})$, l'inégalité $2|x_i y_i| \leq |x_i|^2 + |y_i|^2$ montre que la série $\sum x_i \overline{y_i}$ est absolument convergente et on en déduit que

$$\sum_{i=1}^{\infty} |x_i + y_i|^2 = \sum_{i=1}^{\infty} \left(|x_i|^2 + |y_i|^2 + 2\Re e(x_i \overline{y_i})\right)$$
$$\leq 2 \sum_{i=1}^{\infty} \left(|x_i|^2 + |y_i|^2\right)$$

Ces inégalités montrent que $\mathbf{x} + \mathbf{y}$ est dans $\ell^2(\mathbb{N})$, celui-ci est donc un espace vectoriel. On pose, pour $\mathbf{x}$ et $\mathbf{y}$ dans $\ell^2(\mathbb{N})$,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i} \tag{1.1}$$

On vérifie que cela définit bien un produit scalaire sur $\ell^2(\mathbb{N})$ et les inégalités de Cauchy-Schwarz et de Minkowski s'écrivent

$$\left|\sum_{i=1}^{\infty} x_i \overline{y_i}\right| \leq \sum_{i=1}^{\infty} |x_i \overline{y_i}| \leq \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{\infty} |y_i|^2\right)^{\frac{1}{2}}$$
$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=i}^{\infty} |x_i|^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{\infty} |y_i|^2\right)^{\frac{1}{2}}$$

Soit E un espace vectoriel sur $\mathbb{K}$ et $\langle , \rangle$ un produit scalaire sur E ; l'application $x \longmapsto \sqrt{\langle x, x \rangle}$ est une norme et l'application qui au couple (x, y) associe $d(x, y) = \|x - y\|$ est une distance sur E.

Définition 1.1.8. *Un espace de Hilbert est un espace préhilbertien sur $\mathbb{K}$ qui est complet pour la distance définie par $d(x, y) = \|x - y\|$.*

Un espace de Hilbert est donc encore un espace de Banach[5], dont la norme provient d'un produit scalaire. Une telle norme est parfois appelée *norme de la convergence en moyenne quadratique.*

[5]Stefan Banach (1892-1945), mathématicien polonais, a introduit les espaces qui portent son nom et étudié les applications linéaires dans ces espaces(1920-1930). Il est considéré comme l'un des fondateurs de l'analyse fonctionnelle.

Exemple 1.1.9. - L'espace $\ell^2(\mathbb{N})$ muni du produit scalaire usuel est un espace de Hilbert.

Démonstration. On a vu à l'exemple 1.1.7 que $\ell^2(\mathbb{N})$, muni du produit scalaire (1.1) est un espace préhilbertien, il reste à montrer que $\ell^2(\mathbb{N})$ est complet pour la distance associée au produit scalaire usuel. Soit donc $(\mathbf{x}^p)$ une suite de Cauchy dans $\ell^2(\mathbb{N})$, avec

$$\mathbf{x}^p = (x_1^p, x_2^p, \ldots)$$

Par définition $\sum_{n=1}^{\infty} |x_n^p - x_n^q|^2$ tend vers zéro lorsque p et q tendent vers l'infini ; donc a fortiori pour tout n fixé, la suite numérique $(x_n^p), p \in \mathbb{N}$, est une suite de Cauchy, notons x_n sa limite et $\mathbf{x} = (x_n)$ la suite ainsi définie. Soit $\epsilon > 0$, il existe un entier r tel que, pour $p \geq r$ et $q \geq r$, on ait

$$\sum_{n=1}^{\infty} |x_n^p - x_n^q|^2 < \epsilon$$

Pour tout entier m, on aura a fortiori

$$\sum_{n \leq m} |x_n^p - x_n^q|^2 \leq \epsilon$$

comme il s'agit ici d'une somme finie, on peut faire tendre q vers l'infini et on obtient l'inégalité $\sum_{n \leq m} |x_n^p - x_n|^2 \leq \epsilon$. Cela étant pour tout entier m, on en déduit que

$$\sum_{n=1}^{\infty} |x_n^p - x_n|^2 \leq \epsilon$$

Il en résulte que la suite $\mathbf{x} = (x_n)$ appartient à $\ell^2(\mathbb{N})$ et que lorsque p tend vers l'infini, $\mathbf{x}^p$ tend vers $\mathbf{x}$ dans $\ell^2(\mathbb{N})$. □

L'espace $\ell^2(\mathbb{N})$ joue un rôle fondamental dans l'Analyse Hilbertienne, c'est lui qu'on a longtemps appelé " *l'espace de Hilbert* ".

Exemple 1.1.10. - Soit $C[-1, 1]$ l'espace vectoriel des fonctions continues sur $[-1, 1]$ à valeurs complexes. L'application qui à une fonction f dans $C[-1, 1]$ associe

$$\|f\|_\infty = \sup_{|x| \leq 1} |f(x)|$$

est une norme, dite norme de la convergence uniforme, et l'espace $C[-1, 1]$ muni de cette norme est complet.

Considérons maintenant la forme hermitienne qui, à deux fonctions f et g continues sur $[-1, 1]$, associe

$$\langle f, g \rangle = \int_{-1}^{1} f(t)\overline{g(t)}\, dt$$

C'est un produit scalaire sur $C[-1,1]$ et l'application

$$f \mapsto \|f\|_2 = \left(\int_{-1}^{1} |f(t)|^2 \, dt\right)^{\frac{1}{2}}$$

est une norme sur $C[-1,1]$, dite norme de la convergence en moyenne quadratique. Mais l'espace vectoriel $(C[-1,1], \|.\|_2)$ n'est pas complet :

En effet, considérons la suite des fonctions continues f_n définies sur l'intervalle $[-1,1]$ par

$$f_n(x) = \begin{cases} 0, & \text{si } -1 \leq x \leq -1/n\,; \\ nx+1, & \text{si } -1/n \leq x \leq 0\,; \\ 1, & \text{si } \ 0 \leq x \leq 1. \end{cases}$$

Pour n et m dans $\mathbb{N}$, avec $m \leq n$, la fonction $f_m - f_n$ est positive et nulle en dehors de l'intervalle $[-1/m, 0]$ si bien que

$$\int_{-1}^{1} |f_n(x) - f_m(x)|^2 = \int_{-\frac{1}{m}}^{0} |f_n(x) - f_m(x)|^2 \, dx \leq \frac{2}{m}$$

la suite (f_n) est donc une suite de Cauchy pour la norme $\| \ \|_2$. Nous allons voir qu'elle ne peut pas converger vers une fonction continue. En effet, supposons que f soit la fonction continue sur $[-1,1]$, vers laquelle converge la suite (f_n) ; la quantité $\|f - f_n\|_2^2$ est égale à

$$\int_{-1}^{-\frac{1}{n}} |f(x)|^2 + \int_{-\frac{1}{n}}^{0} |nx + 1 - f(x)|^2 \, dx + \int_{0}^{1} |1 - f(x)|^2 \, dx$$

Sur l'intervalle $[-1/n, 0]$, la fonction $nx+1$ est bornée et donc, lorsque n tend vers l'infini, l'intégrale du milieu tend vers 0. Finalement, il reste

$$\int_{-1}^{0} |f(x)|^2 \, dx + \int_{0}^{1} |1 - f(x)|^2 \, dx = 0$$

c'est-à-dire que la limite f est forcément donnée par

$$f(x) = \begin{cases} 0 & \text{si } -1 \leq x \leq 0\,; \\ 1 & \text{si } 0 < x \leq 1. \end{cases}$$

Ceci est en contradiction avec l'hypothèse de continuité de f. Donc aucune fonction continue ne peut être la limite de (f_n).(voir aussi l'exercice 6).

Exemple 1.1.11. - Soit X un ensemble, Ω une σ-algèbre de X et μ une mesure positive sur Ω. Si f et g sont dans $L^2(X, \Omega, \mu)$, l'inégalité de Hölder montre que $f\overline{g}$ est dans $L^1(X, \Omega, \mu)$ et que l'application

$$(f, g) \mapsto \langle f, g \rangle = \int_X f\overline{g} \, d\mu$$

est un produit scalaire sur $L^2(X,\Omega,\mu)$. Le théorème de Fischer-Riesz assure que $L^2(X,\Omega,\mu)$ muni de ce produit scalaire est complet, c'est donc un espace de Hilbert. Comme cas particulier important de cette situation, on distingue l'espace $L^2(\mathbb{T})$, où $\mathbb{T}$ est l'ensemble des nombres complexes de module 1 (c'est le cercle unité du point de vue ensemble) :

$$L^2(\mathbb{T}) = \left\{ g \mid \|g\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(e^{i\theta})|^2\, d\theta < \infty \right\}$$

Rappelons qu'à tout élément g de $L^2(\mathbb{T})$, on associe la suite donnée, pour n dans $\mathbb{Z}$, par

$$\hat{g}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(e^{i\theta})e^{-in\theta}\, d\theta$$

C'est la suite des coefficients de Fourier[6] de g et le théorème de Parseval se traduit par l'égalité

$$\|g\|_2^2 = \sum_{n\in\mathbb{Z}} |\hat{g}(n)|^2$$

Ainsi, l'application qui à un élément g de $L^2(\mathbb{T})$ associe la suite de ses coefficients de Fourier $(\hat{g}(n))$, $n \in \mathbb{Z}$, est un isomorphisme isométrique de l'espace $L^2(\mathbb{T})$ sur l'espace $\ell^2(\mathbb{Z})$.

Complété d'un espace préhilbertien. Soit E un espace préhilbertien muni du produit scalaire $\langle , \rangle$; soit $\tilde{E}$ son complété relativement à la métrique induite par ce produit scalaire. Il est facile de voir que le produit scalaire $\langle , \rangle$ se prolonge de façon unique en un produit scalaire sur $\tilde{E}$ qui en fait un espace de Hilbert. Autrement dit, tout espace préhilbertien admet un plus petit espace de Hilbert le contenant, à savoir son complété en tant qu'espace vectoriel normé. Ce résultat reste abstrait et la détermination pratique du complété est un problème important pour beaucoup de questions d'Analyse et est à la base de la théorie des espaces de Sobolev par exemple.

L'espace $C[-1,1]$ des fonctions continues sur $[-1,1]$ à valeurs complexes, muni du produit scalaire

$$\langle f, g\rangle = \int_{-1}^{1} f(t)\overline{g(t)}\, dt$$

est un espace préhilbertien qui n'est pas complet pour la norme induite par ce produit scalaire (voir exemple 1.10), son complété est par définition l'espace de Hilbert $L^2((-1,1),dx)$.

[6]Le mathématicien français Jean-Baptiste Joseph FOURIER (1768-1830) est aussi connu comme égyptologue et administrateur. Il a été professeur d'Analyse à l'Ecole Polytechnique de 1796 à 1798. Son étude des séries trigonométriques et ses recherches sur la théorie de la conduction de la chaleur ont eu un impact considérable sur l'évolution actuelle de la Physique Mathématique.

Sous-espaces de Hilbert. Si F est un sous-espace vectoriel d'un espace préhilbertien E, on peut évidemment le munir d'un produit scalaire, restriction de celui de E et F devient préhilbertien ; si de plus E est un espace de Hilbert et F est fermé dans E, F est complet, c'est donc un espace de Hilbert ; on dit alors que F est un sous-espace de Hilbert de E.

Exemple 1.1.12. - Prenons $E = L^2(X, \Omega, \mu)$ et soit Y un sous-ensemble μ-mesurable de X ; l'ensemble des (classes de) fonctions de E qui sont nulles μ-presque partout sur $X \setminus Y$ est un sous-espace vectoriel de E ; il est fermé, car si une suite (f_n) d'éléments de F converge dans E vers un élément f, on a

$$\int_{X\setminus Y} |f|^2 d\mu = \int_{X\setminus Y} |f - f_n|^2 d\mu \leq \int_X |f - f_n|^2 d\mu \to 0, \ (n \to \infty)$$

donc $f = 0$ presque partout sur $X \setminus Y$. De plus F s'identifie naturellement à $L^2(Y, \mu|_Y)$.

Exemple 1.1.13. - Soient n_0 un entier naturel et F_0 l'ensemble des suites $\mathbf{x} = (x_n)$ telles que $x_n = 0$ pour $n > n_0$. F_0 est un sous-espace de Hilbert de l'espace $\ell^2(\mathbb{N})$.

L'ensemble F_1 des éléments $\mathbf{x} = (x_n)$ de $\ell^2(\mathbb{N})$ tels que $x_n = 0$ pour tout entier n, $n \leq n_0$, est aussi un sous-espace de Hilbert de $\ell^2(\mathbb{N})$.

Exemple 1.1.14. - Dans l'espace de Hilbert $L^2(0, 2\pi)$, considérons l'ensemble F des éléments f dont les coefficients de Fourier, $c_n(f)$, sont nuls pour tout entier n strictement négatif. F est, de façon évidente, un sous-espace vectoriel de $L^2(0, 2\pi)$. Pour montrer qu'il est fermé, soit (f_k) une suite de Cauchy d'éléments de F et soit f sa limite, il s'agit de montrer que $c_n(f) = 0$ pour tout $n < 0$. Or, l'égalité de Parseval dit que

$$\|f - f_k\|_2^2 = \sum_{n\in\mathbb{Z}} |c_n(f) - c_n(f_k)|^2$$

Comme le premier membre tend vers 0 lorsque k tend vers l'infini, a fortiori chaque terme du second membre tend vers 0 lorsque k tend vers l'infini, c'est-à-dire que $\lim_{k\to\infty} |c_n(f) - c_n(f_k)| = 0$, pour tout $n \in \mathbb{Z}$. Mais $c_n(f_k) = 0$, pour les entiers n strictement négatifs, il en résulte que $c_n(f) = 0$, pour tout $n < 0$. ☐

EXERCICES

1. On munit $\mathbb{R}^n$ du produit scalaire usuel. En utilisant l'inégalité de Cauchy-Schwarz, montrer que

$$(|a_1| + |a_2| + \cdots + |a_n|)^2 \leq n(a_1^2 + a_2^2 + \cdots + a_n^2)$$

Solution : L'inégalité de Cauchy-Schwarz dit que, pour x et y dans $\mathbb{R}^n$, on a $|\langle x, y\rangle| \leq ||x||\ ||y||$; en particulier, pour $x = (1, 1, \ldots, 1)$ et $y = (|a_1|, \ldots, |a_n|)$, il vient

$$\sum_{i=1}^{n} |a_i| \leq \Big(\sum_{i=1}^{n}(1)\Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{n} a_i^2\Big)^{\frac{1}{2}}$$

L'inégalité voulue s'en déduit. □

2. Autre produit scalaire sur $\mathbb{R}^n$: Soit $A = (a_{ij})$ une matrice carrée ($n \times n$) à coefficients réels. On suppose que A est une matrice symétrique et définie positive c'est-à-dire que, pour tout $\mathbf{x} = (x_1, \ldots x_n) \neq 0$,

$$\langle A\mathbf{x}, \mathbf{x}\rangle = \sum_{i,j=1}^{n} a_{ij}x_ix_j > 0$$

Montrer que l'application $\phi : (\mathbf{x}, \mathbf{y}) \mapsto \langle A\mathbf{x}, \mathbf{y}\rangle$ est un produit scalaire sur $\mathbb{R}^n$. Montrer que tout produit scalaire sur $\mathbb{R}^n$ s'écrit sous la forme précédente.

Solution : Il est facile de vérifier que l'application ϕ est une forme bilinéaire sur $\mathbb{R}^n \times \mathbb{R}^n$. Elle est symétrique si et seulement si $a_{ij} = a_{ji}$ pour tout couple d'indices i, j, c'est-à-dire si la matrice A est symétrique. Si de plus A est définie positive, la forme ϕ est positive et non dégénérée, donc un produit scalaire.
Inversement, soit ψ un produit scalaire sur $\mathbb{R}^n$ et soit (e_i), $1 \leq i \leq n$ la base canonique de $\mathbb{R}^n$. Puisque ψ est bilinéaire, pour $\mathbf{x} = \sum_{i=1}^{n} x_ie_i$ et $\mathbf{y} = \sum_{i=1}^{n} y_je_j$, on a

$$\begin{aligned}\psi(\mathbf{x}, \mathbf{y}) &= \psi\Big(\sum_{i=1}^{n} x_ie_i, \sum_{i=1}^{n} y_je_j\Big) \\ &= \sum_{i,j=1}^{n} \psi(e_i, e_j)x_iy_j \\ &= \langle A\mathbf{x}, \mathbf{y}\rangle\end{aligned}$$

où A est la matrice de coefficient $a_{ij} = \psi(e_i, e_j)$. □

3. Montrer que si B est une forme sesquilinéaire sur un espace vectoriel complexe E, telle que $B(x, x) \in \mathbb{R}$ pour tout $x \in E$, alors B est une forme hermitienne.

Solution : Par hypothèse, $B(x + y, x + y)$ et $B(x + iy, x + iy)$ sont des réels. Comme B est sesquilinéaire, on en déduit que $B(x, y) + B(y, x)$ est réel et $B(y, x) - B(x, y)$ est imaginaire pure. Il en résulte facilement que $\overline{B(x, y)} = B(y, x)$. □
Il faut bien noter que la conclusion n'est pas vraie si E est un espace de Hilbert sur $\mathbb{R}$, donner un exemple.

4. Soit $(E, \langle , \rangle)$ un espace de Hilbert. Montrer que l'application qui, à deux élements x et y associe leur produit scalaire est une fonction continue sur $E \times E$.

Solution : La continuité est une conséquence de l'inégalité de Cauchy-Schwarz : $|\langle x, y\rangle| \leq \|x\|\|y\|$. □

5. Soit E un espace vectoriel sur $\mathbb{C}$, muni d'une norme qui vérifie l'identité du parallélogramme, à savoir

$$\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$$

Pour x et y dans E, on pose

$$\phi(x,y) = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\Big)$$

(a) Montrer que $\phi(x,y) = \overline{\phi(y,x)}$ et que $\phi(x,x) = \|x\|^2$.
(b) Montrer que $\phi(x+z,y) = \phi(x,y) + \phi(z,y)$
(c) Montrer que $\phi(\lambda x, y) = \lambda\phi(x,y)$, pour tout $\lambda \in \mathbb{C}$.
(d) En déduire que toute norme qui satisfait l'identité du parallélogramme est induite par un produit scalaire.

Indication : Le point important est (c). Pour le démontrer , on commence par prouver l'égalité pour $\lambda \in \mathbb{Z}$, puis pour $\lambda \in \mathbb{Q}$. On en déduit l'égalité pour $\lambda \in \mathbb{R}$ et on termine en remarquant que l'égalité est satisfaite pour $\lambda = i$.)

6. Soit a un réel strictement positif et $L^2(-a,a)$ l'espace de Hilbert des (classes de) fonctions mesurables et de carré intégrable sur $(-a,a)$ pour la mesure de Lebesgue. On désigne par $\mathcal{F}$ l'ensemble des éléments de $L^2(-a,a)$ qui sont pairs et par $\mathcal{G}$ celui des éléments impairs. Montrer que $\mathcal{F}$ et $\mathcal{G}$ sont des sous-espaces fermés de $L^2(-a,a)$.

Solution : Pour unélément f de $L^2(-a,a), dx)$, on désigne par $\check{f}$ la fonction définie, pour presque tout x de $(-a,a)$, par : $\check{f}(x) = f(-x)$ et on pose

$$\varphi(f) = f + \check{f}, \quad \psi(f) = f - \check{f}, \quad \varphi_1(f) = f, \quad \varphi_2(f) = \check{f}.$$

Les applications φ_1 et φ_2 sont linéaires et continues de E dans lui-même. C'est donc le cas des applications φ et ψ. Comme $\varphi^{-1}(\{0\}) = \mathcal{G}$ et $\psi^{-1}(\{0\}) = \mathcal{F}$, il en résulte que $\mathcal{F}$ et $\mathcal{G}$ sont deux sous-espaces vectoriels fermés de E. □

7. *Version intégrale de l'inégalité de Minkowski.*
Démontrer l'inégalité suivante

$$\left[\int\left|\int f(x,y)\,dx\right|^2 dy\right]^{\frac{1}{2}} \leq \int\left[\int |f(x,y)|^2\,dy\right]^{\frac{1}{2}} dx$$

Montrer que, plus généralement, pour tout q, avec $1 < q < \infty$,

$$\left[\int \left| \int f(x,y)\,dx \right|^q dy\right]^{\frac{1}{q}} \leq \int \left[\int |f(x,y)|^q\,dy\right]^{\frac{1}{q}} dx$$

Solution : Si l'on désigne par A le premier membre, on voit que A^2 est majoré par

$$\int \left(\int |f(x,y)|\,dx\right)^2 dy = \int \left[\int |f(x,y)| dx\right]\left[\int |f(z,y)| dz\right] dy$$

Intégrant d'abord par rapport à y et utilisant l'inégalité de Hölder, il vient

$$A^2 \leq \int\int \left(\int |f(x,y)|^2 dy\right)^{\frac{1}{2}} \left(\int |f(z,y)|^2 dy\right)^{\frac{1}{2}} dx dz$$

Le second membre est exactement le carré du second membre de l'inéga-lité annoncée. $\square$

1.2 Projection Orthogonale

Soit E un espace métrique, $x \in E$ et F une partie non vide de E.

• Existe-t-il un point $a \in F$ qui soit *le plus proche* de x ? c'est-à-dire tel que $\forall\, y \in F,\ \ d(x,a) \leq d(x,y)$, autrement dit tel que

$$d(x,a) = \inf_{y \in F} d(x,y) \quad \text{ou encore} \quad d(x;F) = d(x,a)$$

• Si un tel point existe, est-il unique ?

Un tel point, quand il existe, est appelé une projection de x sur F.

Si F n'est pas fermé, on peut citer des exemples triviaux où la réponse à la première question est déjà négative (prendre par exemple $E = \mathbb{R}$, $F = [0,1[$ et $x = 2$). Aussi supposerons-nous dans la suite que F est une partie fermée non vide. En général, un point n'a pas forcément de projection sur un sous-ensemble fermé, ou peut en avoir plusieurs (prendre F un cercle et x son centre). Cependant, ce problème a une solution satisfaisante lorsque F est un sous-espace fermé d'un espace de Hilbert E. Cela s'accomplit grâce à la notion d'orthogonalité que nous introduisons maintenant.

Définition 1.2.1. *Deux éléments x et y d'un espace de Hilbert E sont dits orthogonaux si $\langle x, y\rangle = 0$. on écrit alors $x \perp y$.*
On dit que deux parties F et G de E sont orthogonales si tout élément de F est orthogonal à tout élément de G, on écrit alors $F \perp G$.
L'orthogonal d'une partie F de E, noté $F^\perp$, est l'ensemble des éléments de E orthogonaux à F.

Proposition 1.2.2. *Soit E un espace de Hilbert.*

(1) *L'orthogonal d'un sous-ensemble F de E est un sous-espace fermé de E et on a*

$$F \cap F^\perp = \{0\}, \quad F^\perp = (\overline{F})^\perp \quad \textit{et} \quad F \subset (F^\perp)^\perp$$

(2) *Si deux sous-ensembles F et G de E vérifient $F \subset G$, alors leurs orthogonaux vérifient $G^\perp \subset F^\perp$.*

Démonstration. La propriété (2) est triviale. Par ailleurs, il est clair que $F^\perp$ est un sous-espace vectoriel de E, car si x et y sont dans $F^\perp$ alors la linéarité montre que, pour tout z dans F, $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle = 0$. Montrons que $F^\perp$ est fermé : soit (x_n) une suite de Cauchy d'éléments de $F^\perp$ et soit x sa limite. Pour tout y dans F, on a

$$|\langle x, y\rangle| = |\langle x - x_n, y\rangle| \leq \|x - x_n\| \; \|y\|$$

Le membre de droite de cette inégalité tend vers zéro quand n tend vers l'infini et donc $\langle x, y\rangle = 0$. Ceci étant pour tout $y \in F$, on en déduit que x appartient à $F^\perp$ et donc $F^\perp$ est un fermé.

D'autre part, puisque $F \subset \overline{F}$ on a l'inclusion $\overline{F}^\perp \subset F^\perp$. Soit alors x un élément de $F^\perp$ et y un élément de $\overline{F}$, il existe une suite (y_n) d'éléments de F telle que $\lim_{n\to\infty} \|y - y_n\| = 0$. On a alors

$$|\langle x, y\rangle| = |\langle x, y - y_n\rangle| \leq \|x\| \; \|y - y_n\|$$

Le membre de droite tend vers zéro quand n tend vers l'infini et il en résulte que $\langle x, y\rangle = 0$. Ainsi $F^\perp$ est inclus dans $(\overline{F})^\perp$. □

Théorème 1.2.3. (Pythagore[7]) *Si $x_1, x_2, \ldots, x_n$ sont des éléments de E, deux à deux orthogonaux, alors*

$$\|x_1 + x_2 + \cdots x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2$$

Démonstration. Si x_1 et x_2 sont orthogonaux,

$$\|x_1 + x_2\|^2 = \langle x_1 + x_2, x_1 + x_2\rangle = \|x_1\|^2 + 2\Re e(\langle x_1, x_2\rangle) + \|x_2\|^2$$

Comme $\langle x_1, x_2\rangle = 0$, on en déduit le théorème lorsque $n = 2$. Le cas général s'en déduit par induction. □

[7] né vers 580av. J-C et mort vers 490av. J-C, était un mathématicien, philosophe et astronome de la Grèce antique.

Proposition 1.2.4. *Si F et G sont deux sous-espaces fermés d'un espace de Hilbert E et s'ils sont orthogonaux, $F \perp G$, alors l'ensemble $F + G$ des éléments de la forme $x + y$, avec $x \in F$ et $y \in G$, est un sous-espace fermé de E.*

Démonstration. Il est clair que $F + G$ est un sous-espace vectoriel de E. Soit (z_n) une suite de Cauchy dans $F + G$, pour tout n il existe $x_n \in F$ et $y_n \in G$ tels que $z_n = x_n + y_n$, et on a

$$\|z_n - z_m\|^2 = \|x_n - x_m\|^2 + \|y_n - y_m\|^2$$
$$\|x_n - x_m\|^2 \leq \|z_n - z_m\|^2$$
$$\|y_n - y_m\|^2 \leq \|z_n - z_m\|^2$$

donc (x_n) et (y_n) sont des suites de Cauchy dans F et G respectivement. Comme F et G sont fermés, (x_n) converge vers un élément x de F et (y_n) converge vers un élément y de G et par suite $\lim_{n\to\infty} z_n = x + y$ appartient bien à $F + G$. □

La norme d'un espace de Hilbert vérifie les propriétés caractéristiques suivantes :

Théorème 1.2.5. *Dans un espace de Hilbert E la norme vérifie les deux propriétés équivalentes suivantes, valables pour tout x, a et b dans E,*

(1) *L'identité de la médiane :*

$$\|x - a\|^2 + \|x - b\|^2 = 2\|x - \frac{1}{2}(a + b)\|^2 + \frac{1}{2}\|a - b\|^2$$

(2) *L'identité du parallélogramme :*

$$\|x + a\|^2 + \|x - a\|^2 = 2\Big[\|x\|^2 + \|a\|^2\Big]$$

Démonstration. En prenant $b = -a$ dans l'identité de la médiane, on retrouve l'identité du parallélogramme.
Inversement, en appliquant l'identité du parallélogramme aux deux éléments $u = x - a$ et $v = x - b$, on obtient l'identité de la médiane. Il suffit donc de démontrer l'identité du parallélogramme, ce qui se fait facilement en développant son premier membre. □

En interprétant $\|x\|$ comme la longueur du vecteur x, l'identité du parallélogramme traduit la propriété bien connue en géométrie plane qui dit que "*dans un parallélogramme, la somme des carrés des diagonales est égale à la somme des carrés des côtés*".

Ensembles Convexes. Rappelons qu'une partie F d'un espace vectoriel E est dite convexe si elle possède la propriété géométrique suivante : pour

tout a et tout b appartenant à F et pour tout $\alpha \in [0,1]$, l'élément défini par $x_\alpha = (1-\alpha)a + \alpha b$ appartient aussi à F. Quand α décrit l'intervalle $[0,1]$, x_α décrit le "segment de droite" dans E qui joint l'élément a à l'élément b. La convexité exige donc que F contienne les segments joignant deux de ses points.

Les convexes dans $\mathbb{R}$ sont les intervalles. Tout sous-espace vectoriel de E est convexe. De même dans un espace vectoriel normé, la boule fermée $\overline{B}(a,r) = \{x \mid \|x-a\| \leq r\}$ est convexe.

Théorème 1.2.6. (*projection sur une partie convexe complète*). *Soient E un espace préhilbertien et F une partie non vide de E, convexe et complète (par exemple convexe fermé si E est un espace de Hilbert). Pour tout élément x de E, il existe un unique élément a de F tel que $\|x-a\| = d(x;F)$. De plus, a est caractérisé par la propriété suivante :*

$$\Re e\,(\langle x-a, b-a\rangle) \leq 0, \quad \text{pour tout } b \in F$$

Démonstration. 1) Démontrons l'existence d'un élément $a \in F$ vérifiant $\|x-a\| = d(x;F)$. Par définition de la borne inférieure, il existe une suite (a_n) d'éléments de F telle que $\|x-a_n\|$ tend vers $d(x;F)$ lorsque n tend vers l'infini. L'identité de la médiane donne

$$\frac{1}{2}\|a_p - a_q\|^2 = \|x-a_p\|^2 + \|x-a_q\|^2 - 2\|x - \frac{1}{2}(a_p+a_q)\|^2$$

comme F est convexe, $\frac{1}{2}(a_p+a_q) \in F$ et par suite

$$\|x - \frac{1}{2}(a_p+a_q)\|^2 \geq d^2(x;F)$$

On en déduit que

$$\frac{1}{2}\|a_p - a_q\|^2 \leq \|x-a_p\|^2 + \|x-a_q\|^2 - 2d^2(x;F)$$

Le second membre de cette inégalité tend vers 0, lorsque p et q tendent vers l'infini, et il en sera alors de même de $\|a_p - a_q\|^2$. La suite (a_n) est donc une suite de Cauchy ; comme F est supposé complet, elle converge vers un élément a de F et on a

$$\|x-a\| = \lim_{n\to\infty} \|x-a_n\| = d(x;F)$$

2) Démontrons l'unicité de l'élément a. Supposons qu'un autre élément $a' \in F$ réalise aussi l'égalité $\|x-a'\| = d(x;F)$; l'élément $a'' = \frac{1}{2}(a+a')$ appartient à F et par suite $\|x-a''\| \geq d(x;F)$; mais on a aussi

$$\begin{aligned}\|x-a''\|^2 &= \frac{1}{2}\big(\|x-a\|^2 + \|x-a'\|^2\big) - \frac{1}{4}\|a-a'\|^2 \\ &= d^2(x;F) - \frac{1}{4}\|a-a'\|^2\end{aligned}$$

L'égalité $a = a'$ s'en déduit.

3) Montrons maintenant que l'élément a de F est caractérisé par la propriété

$$\Re e\left(\langle x-a, b-a\rangle\right) \leq 0, \quad \forall b \in F$$

Pour tout $\alpha \in [0,1]$, le point $a + \alpha(b-a)$ appartient à F, donc

$$\begin{aligned} \|x-a\|^2 &\leq \|(x-a) - \alpha(b-a)\|^2 \\ &\leq \|x-a\|^2 + \alpha^2\|b-a\|^2 - 2\alpha\Re e\left(\langle x-a, b-a\rangle\right) \end{aligned}$$

la relation cherchée résulte alors de l'arbitraire de α.
Inversement, la propriété $\Re e\left(\langle x-a, b-a\rangle\right) \leq 0$, $\forall b \in F$, implique

$$\begin{aligned} \|x-b\|^2 &= \|(x-a) - (b-a)\|^2 \\ &= \|x-a\|^2 + \|b-a\|^2 - 2\Re e\left(\langle x-a, b-a\rangle\right) \\ &\geq \|x-a\|^2 \end{aligned}$$

Cette inégalité étant vraie pour tout b dans F, il vient $\|x-a\| = d(x;F)$. □

Théorème 1.2.7. (*projection sur un sous-espace de Hilbert*). *Soient E un espace de Hilbert, F un sous-espace de Hilbert et x un élément de E. Il existe un unique élément $a_x \in F$ tel que*

$$\|x - a_x\| = d(x;F)$$

De plus, a_x est l'unique élément de F tel que $x - a_x$ soit orthogonal à F ; il est noté $P_F(x)$ et est appelé la projection orthogonale de x sur F.

Démonstration. La première assertion est déjà démontrée. Pour la seconde, on remarque que si $b \in F$, alors $a + \alpha b \in F$ pour tout complexe α ; le théorème 1.2.6 implique

$$\Re e(\langle x-a, \alpha b\rangle) = \Re e(\langle x-a, \alpha b + a - a\rangle) \leq 0$$

Il suffit ensuite de prendre $\alpha = \pm 1$ et $\alpha = \pm i$.
Réciproquement, soit $a \in F$ tel que $x - a \perp F$. Pour tout $b \in F$, on a

$$\|x-b\|^2 = \|x-a\|^2 + \|b-a\|^2 \geq \|x-a\|^2$$

Il en résulte que $a = P_F(x)$. □

Remarque 1.2.8. - Il faut bien noter que l'utilisation de l'identité de la médiane a été décisive dans la preuve du théorème 1.2.6, et donc pour la validité du théorème 1.2.7. En fait, dans un espace de Banach, ce théorème n'est pas vrai ; par exemple dans l'espace de Banach $(C[0,1], \|.\|_\infty)$ des

fonctions continues sur $[0, 1]$ muni de la norme de la convergence uniforme, l'ensemble

$$F = \{ f \in C[0,1] \mid f(0) = 0 \quad \text{et} \quad 0 \leq f \leq 1 \}$$

est une partie convexe femée et pour tout $f \in F$, on a

$$\|1 - f\|_\infty = \sup_{0 \leq t \leq 1} |1 - f(t)| = 1$$

donc pour tout $f \in F$, $d(1; F) = \|1 - f\|_\infty = 1$. Ainsi, tous les éléments de F minimisent la distance de 1 à F.

Remarque 1.2.9. - La réciproque de ce théorème, à savoir que si dans un espace hilbertien E, un ensemble F est tel que pour tout $x \in E$ il existe un unique $a \in F$ qui est le plus proche de x, alors F est convexe et fermé, a été avancée depuis longtemps, mais elle n'a été ni prouvée ni réfutée.

Corollaire 1.2.10. *Soient E un espace de Hilbert et F un sous-espace de Hilbert de E. Alors les sous-espaces F et $F^\perp$ sont supplémentaires dans E, c'est-à-dire que : $E = F \oplus F^\perp$*

Démonstration. Le sous-espace $F \cap F^\perp$ est réduit à 0 parce que le produit scalaire est non dégénéré ; de plus, pour tout élément x de E on a évidemment $x = x - P_F(x) + P_F(x)$ et $x - P_F(x) \in F^\perp$. □

Corollaire 1.2.11. *L'application P_F est un opérateur linéaire de E dans F qui vérifie, pour tout x et tout y dans E,*

$$\|P_F x\| \leq \|x\|, \quad \langle P_F x, y\rangle = \langle x, P_F y\rangle \quad \text{et} \quad P_F(P_F x) = P_F x$$

Démonstration. Si $x \in E$, on écrit $x = x - P_F(x) + P_F(x)$ et on utilise le théorème de Pythagore

$$\|x\|^2 = \|x - P_F(x)\|^2 + \|P_F(x)\|^2 \geq \|P_F(x)\|^2$$

De plus, comme pour tout $y \in F$, $P_F(y) = y$, et pour tout $x \in E$, $P_F(x) \in F$, on en déduit que $P_F\big(P_F x\big) = P_F(x)$ pour tout $x \in E$. D'autre part, pour x et y dans E, les éléments $P_F x$ et $y - P_F y$ sont orthogonaux et il en résulte que

$$\langle P_F x, y\rangle = \langle P_F x, P_F y\rangle \quad \text{et} \quad \langle x, P_F y\rangle = \langle P_F x, P_F y\rangle$$

Cela termine la preuve du corollaire. □

Remarque 1.2.12. - On doit noter que même si F est un sous-espace vectoriel fermé dans un espace préhilbertien, son orthogonal ne lui est pas nécessairement supplémentaire comme le montre l'exemple suivant : Soit l'espace vectoriel $C[0,1]$, muni du produit scalaire

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,dt$$

Soit F le sous-espace vectoriel des fonctions nulles sur $[0,1/2]$, alors son orthogonal $F^\perp$ est le sous-espace vectoriel des fonctions nulles sur $[1/2,1]$. La fonction constante 1 ne peut être somme d'une fonction nulle sur $[0,1/2]$ et d'une fonction nulle sur $[1/2,1]$ car elle s'annulerait au point $1/2$.

C'est que l'hypothèse "F fermé" n'est pas la bonne hypothèse quand l'espace E n'est pas complet (ce qui est le cas de l'espace $C[0,1]$ muni de la norme induite par le produit scalaire ci-dessus). La bonne hypothèse est que F soit complet comme c'est précisé dans le théorème 1.2.6.

Corollaire 1.2.13. *Soit E un espace de Hilbert et soit F un sous-espace vectoriel (non nécessairement fermé). F est dense dans E si, et seulement si, $F^\perp = \{0\}$.*

Démonstration. D'après la proposition 1.2.2, on a $F^\perp = \left(\overline{F}\right)^\perp$, le résultat s'en déduit. □

Un problème de Minimisation. Soit $v_1, v_2, \ldots, v_k$ des éléments d'un espace de Hilbert E, linéairement indépendants et soit $x \in E$. Il s'agit de trouver un moyen pour calculer la valeur minimum de la quantité

$$\left\| x - \sum_{j=1}^{k} c_j v_j \right\|$$

lorsque $c_1, c_2, \ldots, c_k$ décrivent $\mathbb{K}$, et de trouver les valeurs correspondantes de $c_1, c_2, \ldots, c_k$.

Soit F l'espace vectoriel engendré par les éléments $v_1, v_2, \ldots, v_k$. C'est un sous-espace fermé de E (puisque sa dimension est finie égale à k). La quantité qu'on cherche à minimiser représente la distance de x à l'élément de F donné par $\sum_1^k c_i v_i$. le théorème 1.2.7 précise que $P_F(x)$ est l'unique élément de F qui rend minimum la quantité ci-dessus, cet élément répond donc à la question posée et il s'agit de le déterminer. $P_F(x)$ s'écrit sous la forme $P_F(x) = \sum_{j=1}^k c_j v_j$ et est aussi caractérisé par le fait que $x - P_F(x) \in F^\perp$, ce qui pourrait alors être utilisé pour obtenir des renseignements sur le calcul des coefficients $c_1, c_2, \ldots, c_k$. Posons pour cela

$$a_{ij} = \langle v_j, v_i\rangle, \quad b_i = \langle x, v_i\rangle$$

La propriété $x-P_F(x) \perp F$ implique que $\langle x-P_F(x), v_i\rangle = 0$ pour $1 \leq i \leq k$, ce qui fournit k équations linéaires dont les inconnues sont $c_1, c_2, \dots, c_k$:

$$\sum_{j=1}^{k} a_{ij}c_j = b_i, \quad \text{où} \quad 1 \leq i \leq k$$

L'existence et l'unicité de $P_F(x)$ implique que le déterminant de la matrice (a_{ij}) n'est pas nul, et les (c_j) se calculent en résolvant le système précédent. Soit maintenant γ la valeur minimale de

$$\left\| x - \sum_{j=1}^{k} c_j v_j \right\|$$

Puisque $x - P_F(x)$ est orthogonal à F, on a $\langle x - P_F(x), P_F(x)\rangle = 0$ et donc

$$\gamma^2 = \langle x - P_F(x), x - P_F(x)\rangle = \langle x, x - P_F(x)\rangle = \langle x, x - \sum_{j=1}^{k} c_j v_j\rangle$$

de sorte que

$$\gamma^2 = \|x\|^2 - \sum_{j=1}^{k} \overline{c_j} b_j$$

et notre problème est ainsi résolu.
Venons en maintenant à un cas particulier : On suppose que les éléments $v_1, v_2, \dots, v_k$ sont deux à deux orthogonaux. Alors

$$a_{ij} = \begin{cases} \|v_i\|^2, & \text{si } i = j\,; \\ 0, & \text{sinon.} \end{cases}$$

et par suite le système linéaire que vérifient les coefficients (c_j) donne, pour tout i, $c_i = b_i/\|v_i\|^2$ et on a alors

$$P_F x = \sum_{j=1}^{k} \langle x, v_j\rangle \frac{v_j}{\|v_j\|^2} \quad \text{et} \quad \gamma^2 = \|x\|^2 - \sum_{j=1}^{k} |\langle x, v_j\rangle|^{2/\|v_j\|^2}$$

Si les v_j, $0 \leq j \leq k$ sont orthonormés, ce qui précède se résume comme suit

Théorème 1.2.14. *Soient $v_1, v_2, \dots, v_k$ des éléments deux à deux orthogonaux et de norme 1, dans un espace de Hilbert E, soit F le sous-espace vectoriel engendré par (v_j) et soit x un élément de E. Alors, quels que soient les scalaires $\lambda_1, \lambda_2, \dots, \lambda_k$, on a*

$$\left\| x - \sum_{j=1}^{k} \langle x, v_j\rangle v_j \right\| \leq \left\| x - \sum_{j=1}^{k} \lambda_j v_j \right\|$$

L'égalité a lieu si et seulement si $\lambda_j = \langle x, v_j \rangle$ *pour* $1 \leq j \leq k$*. La projection orthogonale de* x *sur le sous-espace* F *est*

$$P_F x = \sum_{j=1}^{k} \langle x, v_j \rangle v_j$$

La distance γ *de* x *au sous-espace* F *est donnée par*

$$\gamma^2 = \|x - P_F x\|^2 = \|x\|^2 - \sum_{j=1}^{k} |\langle x, v_j \rangle|^2$$

Ainsi, on a trouvé un algorithme constructif permettant de trouver la projection orthogonale d'un élément x de E sur un sous-espace de dimension finie. Cet algorithme sera généralisé au paragraphe 4.

Exemple 1.2.15. - Soit à calculer

$$\min_{a,b,c} \int_{-1}^{1} |x^3 - ax^2 - bx - c|^2 \, dx$$

Considérons $L^2(-1,1)$ l'espace de Hilbert des (classes de) fonctions de carré sommable sur l'intervalle $(-1,1)$ relativement à la mesure de Lebesgue. Le produit scalaire sur $L^2(-1,1)$ est défini par

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\overline{g(x)} \, dx$$

Soit F le sous-espace vectoriel des polynômes de degré inférieur ou égal à deux à coefficients réels et posons $g(x) = x^3$. F étant un sous-espace vectoriel de dimension finie, engendré par les polynômes

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2$$

est donc un convexe fermé de $L^2(-1,1)$.

Il s'agit de trouver le polynôme $p \in F$ qui minimise la distance du polynôme g au sous-espace F et de calculer cette distance. Posons

$$p = c_0 p_0 + c_1 p_1 + c_2 p_2, \quad b_j = \langle g, p_j \rangle, \quad a_{ij} = \langle p_j, p_i \rangle$$

On vérifie rapidement que $b_0 = b_2 = 0$, $b_1 = 2/5$ et que

$$a_{00} = 2, \quad a_{11} = a_{02} = 2/3, \quad a_{22} = 2/5, \quad a_{01} = a_{12} = 0$$

Il en résulte que $c_0 = c_2 = 0$ et $c_1 = 3/5$; la projection orthogonale du polynôme g sur le sous-espace F est donc

$$P_F(g)(x) = 3x/5 \qquad \text{et} \qquad \|g - P_F(g)\|^2 = 8/(175)$$

On en déduit que

$$\min_{a,b,c} \int_{-1}^{1} |x^3 - ax^2 - bx - c|^2 \, dx = 8/(175)$$

et que ce minimum est réalisé lorsque $a = c = 0$ et $b = 3/5$.
Voyons maintenant l'avantage à disposer dans F d'une base orthogonale : on peut vérifier facilement (voir l'exercice 8) que les polynômes

$$v_0(x) = 1, \quad v_1(x) = x, \quad v_2(x) = 3x^2 - 1$$

forment une base orthogonale du sous-espace F. On en déduit que

$$P_F(g) = \langle g, v_0 \rangle \frac{v_0}{\|v_0\|^2} + \langle g, v_1 \rangle \frac{v_1}{\|v_1\|^2} + \langle g, v_2 \rangle \frac{v_2}{\|v_2\|^2}$$

Des raisons de parité évidentes montrent que $\langle g, v_0 \rangle = \langle g, v_2 \rangle = 0$, et que $\langle g, v_1 \rangle = 2/5$ et $\|v_1\|^2 = 2/3$. Il en résulte que $P_F(g) = (3/5)x$.

EXERCICES

1. Soient F_1 et F_2 deux sous-espaces fermés d'un espace de Hilbert E. On suppose $F_2 \subset F_1$. Montrer que $P_{F_2} \circ P_{F_1} = P_{F_2}$.

Solution : Puisque F_1 et F_2 sont des sous-espaces fermés de l'espace de Hilbert E, on peut appliquer le théorème de la projection orthogonale sur un sous-espace fermé. Pour tout x dans E, on pose

$$x_1 = P_{F_1}(x), \quad x_2^1 = P_{F_2}(x_1) \quad \text{et} \quad x_2 = P_{F_2}(x)$$

Le théorème de la projection affirme que x_1 est l'unique élément de F_1 tel que $x - x_1 \perp F_1$; x_2^1 est l'unique élément de F_2 tel que $x_1 - x_2^1 \perp F_2$ et x_2 est l'unique élément de F_2 tel que $x - x_2 \perp F_2$. Comme on a l'inclusion $F_2 \subset F_1$, $x - x_1$ étant orthogonal à F_1, est aussi orthogonal à F_2. Il en résulte que $(x - x_1) + (x_1 - x_2^1) = x - x_2^1$ est aussi orthogonal à F_2. L'unicité montre alors que $x_2 = x_2^1$, c'est-à-dire que
$P_{F_2}(x) = P_{F_2} \circ P_{F_1}(x)$. □

2. Soit E un espace de Hilbert et soit B la boule unité fermée de E, c'est-à-dire $B = \{x \in E, \|x\| \leq 1\}$. Montrer que B un convexe de E.

Solution : soient x et y des éléments de B. Pour tout réel t, $0 \leq t \leq 1$, on a

$$\|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\| \leq 1$$

C'est-à-dire que $tx + (1-t)y$ appartient à B. □

3. Pour f dans l'espace $C([0,1])$ des fonctions continues sur $[0,1]$ à valeurs complexes on pose

$$\|f\| = |f(0)| + \|f\|_1 = |f(0)| + \int_0^1 |f(t)| dt$$

Montrer que cela définit une norme sur $C([0,1])$. Soit F le sous-espace des fonctions qui s'annulent en 0 et soit f_0 la fonction constante égale à 1. Calculer la distance de f_0 à F. Existe-t-il un élément f de F tel que $\|f_0 - f\| = \text{dist}\,(f_0, F)$?

Solution : Il est facile de voir que $f \mapsto \|f\|$ est une norme sur $C([0,1])$. Soit d la distance associée, $d(f,g) = \|f - g\|$ pour f, $g \in C([0,1])$. Par définition, $d(F, f_0) = \inf_{f\in F} \|f_0 - f\|$; c'est-à-dire

$$d(F, f_0) = \inf_{f\in F}(1 + \int_0^1 |1 - f(t)|dt)$$

Il est alors clair que $d(F, f_0) \geq 1$; nous allons montrer l'égalité $d(F, f_0) = 1$. A cet effet, soit (f_n) la suite de fonctions définies par

$$\forall n \geq 2, \quad f_n(t) = \begin{cases} 1, & \text{si } 1/n \leq t \leq 1\,; \\ nt, & \text{si } 0 \leq t \leq 1/n. \end{cases}$$

Pour tout $n \geq 2$, f_n est dans F et on a

$$d(f_0, f_n) = 1 + \int_0^{\frac{1}{n}} (1 - nt)dt = 1 + \frac{1}{2n}$$

Ce qui montre que $d(f_0, f_n)$ tend vers 1, quand n tend vers l'infini. □

4. Calculer

$$\min_{a,b,c} \int_{-1}^{1} |x^3 - ax^2 - bx - c|^2\,dx$$

et trouver le maximum de l'intégrale $\left|\int_{-1}^{1} x^3 f(x)\,dx\right|$ où f est supposée de norme 1 et soumise aux restrictions suivantes

$$\int_{-1}^{1} f(x)\,dx = \int_{-1}^{1} xf(x)\,dx = \int_{-1}^{1} x^2 f(x)\,dx = 0,$$

Solution : Pour la première partie de l'exercice, il suffit de se reporter à l'exemple 1.2.15. La deuxième partie peut s'énoncer sous la forme suivante : calculer

$$\max_{\substack{\|g\|=1 \\ f\perp F}} \left|\int_{-1}^{1} x^3 f(x)\,dx\right|$$

où F est le sous-espace de $L^2(-1,1)$ formé par les polynômes de degré inférieur ou égal à 2. Posons $g(x) = x^3$, la fonction $g - P_F(g)$ appartient à $L^2(-1,1)$ et est orthogonale au sous-espace F et pour tout $f \perp F$, on a

$$|\langle g, f\rangle| = |\langle g - P_F(g), f\rangle| \leq \|g - P_F(g)\|\|f\|$$

Il en résulte que le maximum cherché est atteint pour $f = c(g - P_F(g))$ où $c = \|g - P_F(g)\|^{-1}$. Le calcul de $P_F(g)$ a déjà été fait (voir exemple 1.2.15), celui de la constante c ne présente pas de difficulté. Finalement on trouve

$$c = \frac{5\sqrt{7}}{2\sqrt{2}}, \quad \text{et} \quad f(x) = \frac{5\sqrt{7}}{2\sqrt{2}}(x^3 - \frac{3}{5}x) \qquad \square$$

5. Calculer

$$\min_{a,b,c} \int_0^{\infty} |x^3 - ax^2 - bx - c|^2 e^{-x}\, dx$$

Solution : On désigne par E l'espace de Hilbert des (classes de) fonctions de carré sommable sur $[0, +\infty[$ pour la mesure de densité e^{-x} par rapport à la mesure de Lebesgue, muni du produit scalaire :

$$\forall f, g \in E, \quad \langle f, g \rangle = \int_0^{\infty} f(x)\overline{g(x)} e^{-x} dx$$

On désigne par F l'ensemble des polynômes sur $\mathbb{R}$, de degré inférieur ou égal à 2. C'est un sous-espace de Hilbert de E, dont la dimension est 3. On désigne par g le monôme défini par $g(x) = x^3$. L'intégrale

$$\int_0^{\infty} (x^3 - ax^2 - bx - c)^2\, e^{-x} dx$$

représente le carré de la distance de l'élément g au sous-espace F. Le théorème de la projection sur un sous-espace de Hilbert assure qu'il existe un unique polynôme $P_F(g)$ de F tel que $g - P_F(g) \perp F$ et tel que

$$\min_{a,b,c} \int_0^{\infty} (x^3 - ax^2 - bx - c)^2\, e^{-x} dx = \|g - P_F(g)\|^2$$

Il s'agit donc de trouver le polynôme $P_F(g)$ et de calculer sa distance à g. Le polynôme $P_F(g)$ s'écrit sous la forme

$$P_F(g)(x) = \alpha x^2 + \beta x + \gamma$$

de plus, comme $g - P_F(g) \in F^{\perp}$, on a les trois équations

$$\langle g - P_F(g), 1 \rangle = 0, \quad \langle g - P_F(g), x \rangle = 0 \quad \text{et} \quad \langle g - P_F(g), x^2 \rangle = 0$$

En utilisant la formule suivante

$$\forall n \in \mathbb{N}, \quad \int_0^{\infty} x^n e^{-x} dx = n!$$

qu'on peut démontrer par récurrence, les trois équations précédentes se traduisent par les suivantes :

$$\begin{cases} 3! - \beta - \gamma = 0 \\ 4! - \alpha 3! - \beta 2! - \gamma = 0 \\ 5! - \alpha 4! - \beta 3! - \gamma 2! = 0 \end{cases}$$

c'est-à-dire

$$\begin{cases} 2\alpha + \beta + \gamma = 6 \\ 6\alpha + 2\beta + \gamma = 24 \\ 24\alpha + 6\beta + 2\gamma = 120 \end{cases}$$

On en déduit facilement que $\alpha = 6$, $\beta = 0$ et $\gamma = -12$ et par suite

$$P_F(g)(x) = 6x^2 + 12$$

Il reste à calculer $\|g - P_F(g)\|$, ce qui ne présente aucune difficulté. □

6. Soit F un sous-espace fermé d'un espace de Hilbert E et soit $a \in E$. Démontrer que

$$\min\{ \|x - a\| : x \in F\} = \max\{|\langle a, y\rangle| : y \in F^{\perp}, \|y\| = 1 \}$$

Solution : Par définition même de la projection orthogonale de a, il est clair que le minimum de $\|x - a\|$, lorsque x varie dans F, est égal à $\|a - P_F a\|$. D'un autre côté, pour tout élément y dans E et de norme 1, $|\langle a - P_F a, y\rangle| \leq \|a - P_F a\|$ et l'égalité a lieu si, et seulement si, y et $a - P_F a$ sont proportionnels. Dans ce cas $y = c(a - P_F a)$ où $c = \|a - P_F a\|^{-1}$; la relation annoncée s'en déduit. □

7. Soit $C[0,1]$ l'espace des fonctions continues sur $[0,1]$, muni de la norme de la convergence uniforme. Soit F l'ensemble des fonctions $f \in C[0,1]$ telles que

$$\int_0^{\frac{1}{2}} f(t)\,dt - \int_{\frac{1}{2}}^1 f(t)\,dt = 1$$

Démontrer que F est un sous-ensemble fermé convexe de $C[0,1]$ qui ne contient pas d'élément de norme minimale.

8. Montrer que les polynômes définis par $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = 3x^2 - 1$ et $p_3(x) = 5x^3 - 3x$ sont deux à deux orthogonaux dans l'espace $L^2((-1,1), dx)$.

Solution : Pour des raisons de parité évidentes, on a

$$\langle p_0, p_1\rangle = \langle p_0, p_3\rangle = \langle p_1, p_2\rangle = \langle p_2, p_3\rangle = 0$$

On vérifie d'autre part que

$$\langle p_0, p_2\rangle == [x^3 - x]_{-1}^1 = 0, \langle p_1, p_3\rangle = \left[x^5 - x^3\right]_{-1}^1 = 0$$

Donc p_0, p_1, p_2 et p_3 sont deux à deux orthogonaux. □

1.3 Dualité et théorème de Représentation de Riesz

La construction de l'espace dual d'un espace vectoriel topologique est une opération importante dans la mesure où cet espace dual possède de bonnes propriétés quand l'espace initial en possède lui-même. Nous allons voir, comme conséquence du théorème de la projection, que tout espace de Hilbert est isomorphe à son dual (topologique).
Rappelons d'abord quelques propriétés élémentaires valables dans le cadre des espaces vectoriels normés.

Définition 1.3.1. *Une forme linéaire continue sur un espace de Hilbert E est une application linéaire $L : E \to \mathbb{K}$ telle que, quelle que soit la suite (x_n) convergente vers x dans E, la suite $(L(x_n))$ converge vers $L(x)$.*

L'ensemble des formes linéaires continues sur E est appelé le dual topolo-gique de E et se note souvent E'.

Proposition 1.3.2. *Une forme linéaire L sur E est continue si, et seulement si, il existe une constante c telle que*

$$|L(x)| \leq c\|x\|, \quad \text{pour tout } x \in E$$

Démonstration. Il est clair que s'il existe c telle que l'on ait $|L(x)| \leq c\|x\|$ pour tout $x \in E$, alors L est continue puisque, si une suite (x_n) converge vers x dans E,

$$|L(x) - L(x_n)| = |L(x - x_n)| \leq c\|x - x_n\| \to 0, \ (n \to \infty)$$

Inversement, si L est continue l'image réciproque par L de la boule ouverte centrée à l'origine et de rayon 1 est un ouvert de E contenant 0, il existe donc $\eta > 0$ tel que $\|x\| \leq \eta$ implique $\|Lx\| < 1$. Soit x un élément arbitraire de E non nul, l'élément $\eta\|x\|^{-1}x$ a une norme égale à η donc $\eta\|x\|^{-1}\|Lx\| < 1$ et par suite, pour tout $x \in E$,

$$\|Lx\| < \eta^{-1}\|x\|$$

Cela montre que la forme linéaire L est bornée par η^{-1}. □

Pour $L \in E'$, on pose

$$\|L\| = \sup_{\|x\| \leq 1} \|Lx\|$$

On vérifie facilement que l'on définit ainsi une norme sur E'. Cette norme se définit encore par la relation

$$\|L\| = inf\left\{c > 0 \,/\, \|Ax\| \leq c\|x\|, \ \forall x \in E\right\}$$

Autrement dit, pour tout x dans E, on a

$$\|L(x)\| \leq \|L\| \, \|x\|$$

et le réel $\|L\|$ est la plus petite constante réalisant cette propriété.

Remarque 1.3.3. - Si E est de dimension finie, par exemple E est $\mathbb{R}^n$ ou $\mathbb{C}^n$, une application linéaire de E dans F est nécessairement continue. On a vu au paragraphe 2 que dans un espace préhilbertien, un opérateur de projection orthogonale est continu.

Cependant, si E n'est pas de dimension finie, il existe des formes linéaires qui ne sont pas continues. Par exemple sur l'espace préhilbertien $C[0,1]$ muni du produit scalaire

$$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\, dx$$

la forme linéaire $\delta : f \mapsto f(0)$ n'est pas continue. En effet la suite de fonctions (f_n) définies par $f_n(x) = (1-x)^n$ vérifie

$$\forall n \in \mathbb{N}, \ \ \delta(f_n) = 1 \quad \text{et} \quad \|f_n\| = (2n+1)^{-\frac{1}{2}} \to 0, \ (n \to \infty)$$

Fixons nous un élément a de E et définissons la forme linéaire L_a sur E par $L_a(x) = \langle x, a\rangle$. L'inégalité de Cauchy-Schwarz permet de montrer que L_a est continue et que $\|L_a\| \leq \|a\|$; en fait on a l'égalité car $L_a(a) = \|a\|^2$.

Théorème 1.3.4. (de Représentation de Riesz[8])

(1) *Soit E un espace de Hilbert et soit L une forme linéaire continue sur E. Alors il existe un unique élément a de E tel que*

$$\forall\, x \in E, \ \ L(x) = \langle x, a\rangle \quad \textit{et} \quad \|L\| = \|a\|$$

(2) *Inversement, tout élément a de E définit une forme linéaire continue L_a sur E par la formule*

$$L_a(x) = \langle x, a\rangle, \quad \forall x \in E$$

Ainsi l'application qui à un élément $a \in E$ associe la forme linéaire continue L_a est un isomorphisme (anti-linéaire) topologique de E sur son dual E'

Démonstration. Il suffit de considérer le cas d'une forme linéaire continue L, non identiquement nulle. Soit alors $F = \ker(L)$, c'est un hyperplan fermé de E (car L est continue), qui est distinct de E et le corollaire 1.2.10 nous

[8] Frédéric RIESZ (1880-1956), mathématicien d'origine hongroise, exerça une influence profonde sur le développement des mathématiques modernes. Il est, avec S. Banach, l'un des principaux fondateurs de l'analyse fonctionnelle.

dit que $E = F \oplus F^\perp$. Soit y_0 un élément non nul de $F^\perp$, alors $L(y_0) \neq 0$. Pour tout $x \in E$, on pose

$$u = x - \frac{L(x)}{L(y_0)} y_0$$

Il est clair que $L(u) = 0$, c'est-à-dire que u appartient à F, et comme y_0 appartient à $F^\perp$, on a $\langle u, y_0 \rangle = 0$. Cela s'écrit, en remplaçant u par son expression, $\langle x, y_0 \rangle - \|y_0\|^2 L(x)/L(y_0) = 0$. On en déduit que

$$L(x) = \langle x, y_0 \rangle \frac{L(y_0)}{\|y_0\|^2}, \quad \forall\, x \in E$$

et il suffit alors de poser $a = \|y_0\|^{-2}\overline{L(y_0)}y_0$. Il est facile de voir que l'élément a est unique. □

Corollaire 1.3.5. *Soit* (X, Ω, μ) *un espace mesuré et soit* L *une forme linéaire continue sur l'espace* $L^2(X, \Omega, \mu)$, *alors il existe un unique élément* $f_0 \in L^2(\mu)$ *tel que*

$$L(g) = \int_X g\overline{f_0}\, d\mu.$$

Remarque 1.3.6. - Le théorème de représentation de Riesz nous permet donc toujours d'identifier les espaces E et E'. Cependant il faut prendre garde de ne pas identifier "simultanément" les espaces duaux de deux espaces de Hilbert dont l'un est contenu dans l'autre. Plus précisément soient E_0 et E_1 deux espaces de Hilbert distincts dont les normes respectives sont notées $\| \ \|_0$ et $\| \ \|_1$. On suppose que

$$E_1 \subset E_0 \quad \text{et} \quad \exists\, c > 0 \mid \forall\, u \in E_1,\ \|u\|_0 \leq \|u\|_1$$

(c'est-à-dire que l'injection de E_1 dans E_0 est continue) et que de plus E_1 est dense dans E_0.

On peut donc identifier E_0', dual de E_0, à une partie de E_1' dual de E_1 et écrire $E_0' \subset E_1'$, avec injection continue (car E_1 est dense dans E_0). Il est alors clair que si on décide d'identifier par le théorème de Riesz E_0 et E_0', on ne peut plus à la fois considérer E_1 comme un sous-espace vectoriel de E_0 et identifier E_1 avec E_1'. Souvent, lorsque E_0 est l'espace fonctionnel $L^2(X, \Omega, \mu)$, on décide d'identifier E_0 et E_0', et de ce fait on s'interdit d'identifier un espace de Hilbert plus petit (c'est-à-dire contenu dans $L^2(X, \Omega, \mu)$) à son dual.

Topologie faible. L'espace dual E' d'un espace vectoriel normé E permet d'introduire une nouvelle topologie sur E, appelée *la topologie faible* qui est un bon exemple de topologie définie par une famille de semi-normes. Cette topologie prend une grande importance dans le cas des espaces de Hilbert

grâce à l'identification possible, par le théorème de Riesz, de l'espace et de son dual topologique.

Rappelons que, dans un espace de Banach E, une suite (x_n) converge faiblement vers x si pour toute forme linéaire L, continue sur E, la suite $(L(x_n))$ converge vers $L(x)$. Si E est un espace de Hilbert, la convergence faible se traduit, grâce au théorème de représentation de Riesz, par la définition suivante.

Définition 1.3.7. *Une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments d'un espace de Hilbert E converge faiblement vers a si pour tout élément $y \in E$,*

$$\lim_{n\to+\infty} \langle x_n, y\rangle = \langle a, y\rangle.$$

Commentaire. Notons bien que si, pour tout $y \in E$, la suite $(\langle x_n, y\rangle)$ converge, alors la suite (x_n) converge faiblement vers un élément $a \in E$ car, la suite de formes linéaires continues (L_{x_n}) étant convergente en tout point $y \in E$, le théorème de la borne uniforme (voir l'annexe) assure que

(i) La suite (L_{x_n}) est uniformément bornée, c'est-à-dire qu'il existe M telle que $\|x_n\| \leq M$ pour tout $n \in \mathbb{N}$.

(ii) Il existe une forme linéaire continue L telle que, pour tout y dans E, $\lim_{n\to+\infty} \langle x_n, y\rangle = L(y)$.

Le théorème de représentation de Riesz garantit alors l'existence de $a \in E$ tel que $L(y) = \langle a, y\rangle$. Notons aussi que l'assertion (i) dit que *toute suite faiblement convergente est bornée.*

Evidemment si (x_n) converge dans E vers a, alors (x_n) converge faiblement vers a, puisque $|\langle x_n, y\rangle - \langle a, y\rangle| \leq \|x_n - a\| \, \|y\|$. On dira que la convergence en norme entraîne la convergence faible. Mais la réciproque n'est pas vraie, ce que montre l'exemple de la suite de fonctions $e_n(t) = e^{int}, 0 \leq t \leq 2\pi$ emprunté à la théorie des séries de Fourier. La suite e_n converge faiblement vers 0 dans l'espace de Hilbert $L^2(0, 2\pi)$, mais elle ne converge pas en moyenne quadratique puisque, pour $n \neq m$,

$$\|e_n - e_m\|_{L^2(0,2\pi)} = 2\sqrt{\pi}$$

En fait une différence essentielle entre "*convergence en norme*" et "*convergence faible*" de (x_n) vers a est que la première propriété entraîne la relation $\lim_{n\to\infty} \|x_n\| = \|a\|$, tandis que cette relation ne découle pas de la seconde propriété. Mais il y a plus, la convergence faible couplée avec la relation précédente entraîne précisément la convergence en norme :

Théorème 1.3.8. *Dans un espace de Hilbert, une suite (x_n) converge vers a si et seulement si*

(a) *La suite (x_n) converge faiblement vers a*

(b) *La suite (x_n) vérifie $\lim_{n\to\infty} \|x_n\| = \|a\|$.*

Démonstration. Nous avons en effet

$$\|x_n - a\|^2 = \|x_n\|^2 + \|a\|^2 - \langle x_n, a\rangle - \langle a, x_n\rangle$$

Il suffit alors de passer à la limite, quand n tend vers l'infini. □

Une autre différence importante entre la topologie forte et la topologie faible est relative à la notion de compacité. Ce concept, on le sait, est lié à l'un des plus importants théorèmes de l'Analyse élémentaire : le théorème de Bolzano[9]-Weierstrass[10].

Théorème 1.3.9. *Si (x_n) est une suite d'éléments d'un espace de Hilbert E, telle que $\|x_n\| \leq 1$, $\forall n$; alors on peut en extraire une sous-suite qui converge faiblement dans E.*

On exprime ce résultat en disant que "la boule unité fermée de E est faiblement compacte".

Démonstration. Par hypothèse $\|x_n\| \leq 1$. La suite $(\langle x_n, x_1\rangle)$ est donc bornée dans $\mathbb{C}$ et le théorème de Bolzano-Weierstrass assure l'existence d'une sous-suite convergente $(\langle x_{n1}, x_1\rangle)$. La suite $(\langle x_{n1}, x_2\rangle)$ admet à son tour une sous-suite convergente $(\langle x_{n2}, x_2\rangle)$. Répétant cet argument, on en déduit l'existence d'une sous-suite (x_{nj}) de la suite (x_n) telle que

$$(\langle x_{nj}, x_p\rangle) \quad \text{converge pour tout} \quad p \leq j$$

Par le procédé diagonal de Cantor[11] on montre que, pour tout $p \in \mathbb{N}$, la suite $(\langle x_{nn}, x_p\rangle)$ converge. On en déduit que pour tout $y \in E$, la suite $(\langle x_{nn}, y\rangle)$ converge. Cela se vérifie d'abord pour tout élément y du sous-espace fermé F engendré par la famille $\{x_1, x_2, \ldots, x_n, \ldots\}$. Si ce sous-espace coïncide avec E notre assertion se trouve démontrée ; dans le cas contraire, on écrit $E = F \oplus F^\perp$, et tout élément $y \in E$ s'écrit $y = y_1 + y_2$ avec $y_1 \in F$ et $y_2 \in F^\perp$, il s'ensuit que $\langle x_{nn}, y\rangle = \langle x_{nn}, y_1\rangle$. La suite $(\langle x_{nn}, y\rangle)$ est dans ce cas aussi convergente. □

[9]Bernhard Bolzano (1781-1848), est un philosophe et mathématicien tchèque, d'origine italienne. Il est, avec Weierstrass, l'un des créateurs de la théorie des fonctions réelles. Il a laissé une œuvre étendue et importante que ses contemporains ont presque totalement ignorée. Il s'est notamment proposé d'établir une doctrine de l'infini à partir de concepts ensemblistes.

[10] Karl Theodor Wilhelm WEIERSTRASS (1815-1897) est un mathématicien allemand, il a été un des fondateurs de la théorie moderne des fonctions. Il a notamment complété les travaux d'Abel et de Jacobi sur les fonctions elliptiques. Il a eu une influence déterminante sur la Mathématique de la fin du XIXème siècle.

[11] Georg Cantor (1845–1918) né à Saint-Petersbourg, a bouleversé par ses travaux la pensée scientifique du XXème sciècle. Il s'est attaqué aux concepts de l'infini et introduisit à cette occasion les nombres transfinis et les bases de la théorie des ensembles. Hilbert avec son ami Minkowski étaient des partisans de Cantor à une époque où ils étaient fort peu nombreux. A ce propos, Hilbert disait :"il ne nous chasseront pas du paradis que Cantor a crée pour nous".

Ce théorème, dont une illustration est donnée à l'exercice 6, est parfois appelé *théorème de sélection* et, à travers la preuve ci-dessus, on comprend tout à fait la raison. A force d'extraire il risque d'y avoir pénurie, et c'est pour éviter celle-ci que Cantor a inventé cet algorithme de sélection appelé encore procédé diagonal.

Remarque 1.3.10. - Rappelons qu'en contraste avec ce théorème, un autre théorème de Riesz nous apprend que la boule unité d'un espace normé (et donc d'un espace de Hilbert) n'est compacte pour la topologie déduite de la norme (*topologie forte*) que si l'espace est de dimension finie.

EXERCICES

1. Soit L une forme linéaire continue sur un espace de Hilbert E, non identiquement nulle. Montrer que $\dim(\ker(L)^{\perp}) = 1$.

   Solution : L est une forme linéaire continue non identiquement nulle donc ker (L) est un sous espace vectoriel fermé de E, c'est donc un sous espace de Hilbert de E. D'après le corollaire 2.10, on a

$$E = \ker(L) \oplus (\ker(L))^{\perp}$$

   Cela veut dire que tout élément x de E s'écrit de façon unique sous la forme $x = x_0 + x_1$ avec $x_0 \in \ker(L)$ et $x_1 \in (\ker(L))^{\perp}$. Soit y_0 un élément non nul de $\ker(L)^{\perp}$, tout élément x de E peut s'écrire sous la forme :

$$x = \left(x - \frac{L(x)}{L(y_0)}y_0\right) + \frac{L(x)}{L(y_0)}y_0$$

   Il est facile de voir que le premier terme du second membre est dans $\ker(L)$ et que le second terme est dans $\ker(L)^{\perp}$. Il en résulte, en particulier, que $(\ker(L))^{\perp}$ est engendré par y_0 et est donc de dimension 1. □

2. Soit $E = \ell^2(\mathbb{N})$. On se fixe un entier p et on considère la forme linéaire définie sur E par $L(x) = x_p$, où $x = (x_i), i \in \mathbb{N}$. Trouver l'élément $a \in E$ tel que $L = L_a$.

   Réponse : $a = e_p$ le $p^{\text{ème}}$ vecteur de la base canonique de $\ell^2(\mathbb{N})$.

3. Soit $E = \ell^2(\mathbb{N})$.
   (a) Montrer que si $\alpha = (\alpha_n)$ est un élément de E, la série entière $\sum \alpha_n z^n$ a un rayon de convergence supérieur ou égal à 1.
   (b) Soit λ un nombre complexe, $|\lambda| < 1$, et L la forme linéaire définie sur E par $L(\alpha) = \sum \alpha_n \lambda^n$. Trouver l'élément $a \in E$ tel que $L = L_a$
   (c) Quelle est la norme de la forme linéaire L ?

   Réponse : a) Puisque $\alpha = (\alpha_n)$ appartient à $\ell^2(\mathbb{N})$, le terme général α_n tend vers 0 quand n tend vers l'infini. Il en résulte que, pour $|z| < 1$, la série $\sum \alpha_n z^n$ est absolument convergente.

b)–c) On vérifie que $a = (\overline{\lambda}^n)$ et que la norme de L est égale à celle de a, c'est-à-dire

$$\|L\| = \left(\sum_n |\lambda^{2n}|\right)^{\frac{1}{2}} = \frac{1}{1-|\lambda|^2}$$

4. Montrer que dans l'espace de Hilbert $L^2(0, 2\pi)$, la suite définie par $f_n(x) = \sin(nx)$, où $n \in \mathbb{N}$, converge faiblement vers zéro lorsque n tend vers l'infini mais ne converge en norme vers aucune limite.

Indication : Le lemme de Riemann montre que si f est dans $L^2(0, 2\pi)$, alors

$$\lim_{n\to\infty} \int_0^{2\pi} f(x) \sin(nx)\, dx = 0$$

Cela exprime la convergence faible de la suite (f_n) vers 0. D'autre part, il est facile de voir que $\|f_n\| = \sqrt{\pi}$, pour tout entier n ; la suite (f_n) ne converge donc pas au sens de la norme de $L^2(0, 2\pi)$. □

5. Montrer que, dans un espace de Hilbert E, si (x_n) converge faiblement vers x et si (y_n) est une suite qui converge vers y, alors la suite $(\langle x_n, y_n\rangle)$ converge vers $\langle x, y\rangle$.

Solution : On a

$$\begin{aligned}|\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n - y\rangle + \langle x_n - x, y\rangle| \\ &\le |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\le \|y_n - y\|\, \|x_n\| + |\langle x_n - x, y\rangle|\end{aligned}$$

Puisque (x_n) converge faiblement vers x, le deuxième terme du second membre tend vers 0 quad n tend vers l'infini. D'autre part, la suite (x_n) est bornée, car elle est faiblement convergente, et comme $\|y - y_n\|$ tend vers 0 quand n tend vers l'infini, il en sera de même du premier terme du second membre. Donc, $\langle x_n, y_n\rangle$ converge vers $\langle x, y\rangle$. □

6. Soient $E = L^2[0, 1]$ et (f_n), $n \in \mathbb{N}$ une suite bornée d'éléments de E. On pose

$$F_n(x) = \int_0^x f_n(t)\, dt, \quad \text{où} \quad n \in \mathbb{N}$$

Montrer que la suite (f_n) converge faiblement vers 0 si, et seulement si, la suite (F_n) converge uniformément vers 0 sur $[0, 1]$.
En déduire que (f_n) converge faiblement vers un élément f de E si et seulement si la suite (F_n) converge uniformément sur $[0, 1]$ vers la primitive de f qui s'annule en 0.

Solution : Supposons d'abord que (f_n) converge faiblement vers 0 et soit $C > 0$ telle que $\|f_n\| \leq C$, pour tout n. La convergence faible implique en particulier que (F_n) converge simplement vers 0, et l'inégalité de Cauchy-Schwarz montre que si $0 \leq x \leq y \leq 1$, alors $|F_n(y) - F_n(x)| \leq C|y - x|^{\frac{1}{2}}$ et ces deux faits entraînent la convergence uniforme. En effet, pour tout $\epsilon > 0$, soit N fixé tel que $N^{-1} \leq \epsilon^2/(2C^2)$ et $I = \{0, 1/N, 2/N, \ldots, (N-1)/N\}$. La convergence simple implique l'existence d'un entier n_0 tel que pour tout $n \geq n_0$ et tout x dans I, on ait $|F_n(x)| \leq \epsilon/2$. Maintenant, si x est dans $[0, 1]$, il existe un x_0 dans I, tel que $|x - x_0| \leq 1/(2N)$, et par suite pour tout $n \geq n_0$ on a $|F_n(x) - F_n(x_0)| \leq C|x - x_0|^{1/2} \leq C/\sqrt{2N} \leq \epsilon/2$, et finalement $|F_n(x)| \leq \epsilon$ si $n \geq n_0$. En sens inverse, la convergence uniforme sur $[0, 1]$ de (F_n) vers 0 se traduit par $\lim_{n\to\infty}\langle f_n, g\rangle$ où $g = \chi_{[0,x]}$. Par linéarité, on aura cette propriété pour toute fonction g en escalier. Enfin, le théorème de Banach-Steinhaus fournit le cas général. □

1.4 Bases Hilbertiennes

Dans ce paragraphe, on généralise aux espaces de Hilbert les notions de base orthogonale ou orthonormale dans l'espace euclidien. Cela passe inévitablement par la considération de séries de vecteurs. Dans un espace de Banach, un des critères généraux de convergence d'une série, dont on dispose, est celui de la convergence normale (voir l'annexe). Dans le cadre des espaces de Hilbert, et pour les séries d'éléments deux à deux orthogonaux, il est possible d'avoir un autre critère.

Théorème 1.4.1. *Soient E un espace de Hilbert et (x_n) une suite d'éléments de E deux à deux orthogonaux. La série $\sum_i x_i$ converge si et seulement si la série numérique $\sum_i \|x_i\|^2$ converge. Dans ce cas, on a la relation de Pythagore généralisée*

$$\left\|\sum_{i=1}^{\infty} x_i\right\|^2 = \sum_{i=1}^{\infty} \|x_i\|^2$$

Démonstration. Soit (y_n) la suite des sommes partielles de la série $\sum_i x_i$ et (α_n) la suite des sommes partielles de la série $\sum_i \|x_i\|^2$. Le théorème de Pythagore assure que, pour deux entiers quelconques $p < q$,

$$\left\|\sum_{p+1}^{q} x_i\right\|^2 = \sum_{p+1}^{q} \|x_i\|^2$$

ce qui s'écrit $\|y_q - y_p\|^2 = |\alpha_q - \alpha_p|^2$. Ainsi, la suite (y_n) est une suite de Cauchy si et seulement si la suite (α_n) l'est aussi, et dans ce cas on a

$$\lim_{n\to\infty}\left\|\sum_{i=1}^{n} x_i\right\|^2 = \lim_{n\to\infty}\sum_{i=1}^{n}\|x_i\|^2$$

ce qui termine la démonstration. □

On passe maintenant à la généralisation de la notion de base algébrique d'un espace vectoriel dans le contexte des espaces de Hilbert. On rappelle d'abord quelques définitions bien connues.

Définition 1.4.2. *Soient E un espace vectoriel sur $\mathbb{K}$, $x_1, x_2, \ldots, x_k$ des éléments de E et $c_1, c_2, \ldots, c_k$ des scalaires, alors $c_1x_1 + c_2x_2 + \cdots c_kx_k$ est appelé une combinaison linéaire de $x_1, x_2, \ldots, x_k$.*

Soit $\mathcal{F} = \{\, x_i, i \in I\,\}$, une famille d'éléments de E. L'ensemble V constitué de toutes les combinaisons linéaires finies d'éléments de la famille $\mathcal{F}$ est évidemment un sous-espace vectoriel de E, appelé le sous-espace vectoriel engendré par cette famille. C'est le plus petit sous-espace vectoriel de E contenant les éléments x_i, $i \in I$.

Supposons maintenant E un espace de Hilbert et soit $F = \overline{V}$ l'adhérence de V ; c'est un sous-espace complet (car fermé) de E. C'est le plus petit sous-espace de Hilbert contenant la famille $\mathcal{F}$, il est appelé le sous-espace de Hilbert engendré par la famille $\{\, x_i, i \in I\,\}$.

Dans un espace de Hilbert E, une famille $\{\, x_i, i \in I\,\}$ est dite orthogonale si ses éléments sont deux à deux orthogonaux, c'est-à-dire si $\langle x_i, x_j\rangle = 0$ pour tout $i \neq j$. Elle est dite orthonormale (ou orthonormée) si on a de plus $\|x_j\| = 1$ pour tout i.

Le théorème suivant généralise le théorème 1.2.14.

Théorème 1.4.3. *Soit E un espace de Hilbert, soit $\{\, e_i, i \in \mathbb{N}\,\}$ une suite orthonormée d'éléments de E et soit F le sous-espace de Hilbert qu'elle engendre*

(a) *Pour tout $x \in E$, la série vectorielle $\sum_i \langle x, e_i\rangle e_i$ est convergente, sa somme est égale à $P_F(x)$ et on a l'inégalité de Bessel*[12]

$$\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2 = \|P_F(x)\|^2 \leq \|x\|^2$$

[12] L'astronome allemand Friedrich Wilhelm BESSEL (1784-1846) a beaucoup correspondu avec Gauss. Il a contribué à la théorie du Potentiel et son étude des perturbations des trajectoires des planètes l'a conduit à développer les fonctions spéciales qui portent son nom.

(b) *Quels que soient x et y dans E, la série numérique $\sum_i \langle x, e_i\rangle\overline{\langle y, e_i\rangle}$ est absolument convergente et on a*

$$\sum_{i=1}^{\infty}\langle x, e_i\rangle\overline{\langle y, e_i\rangle} = \langle P_F(x), P_F(y)\rangle$$

Démonstration. Pour tout entier $n \geq 1$, on désigne par F_n le sous-espace vectoriel engendré par les éléments $\{e_1, \dots e_n\}$ dont ils forment par hypothèse une base orthonormée. Il suit du théorème 1.2.14 que, pour tout élément x de E,

$$P_{F_n}(x) = \sum_{i=1}^{n}\langle x, e_i\rangle e_i \quad \text{et} \quad \|P_{F_n}(x)\|^2 = \sum_{i=1}^{n}|\langle x, e_i\rangle|^2$$

L'inégalité $\|P_{F_n}(x)\| \leq \|x\|$ assure la convergence de la série $\sum_i |\langle x, e_i\rangle|^2$ et, tenant compte du théorème 1.4.1, celle de la série $\sum_i\langle x, e_i\rangle e_i$. La somme de cette dernière est un élément a de F qui vérifie par construction

$$\langle x - a, e_i\rangle = 0, \quad \text{pour} \quad i \geq 1$$

Mais nous savons bien que cela est une propriété caractéristique de $P_F(x)$. On a donc $a = P_F(x)$ et de nouveau le théorème 1.4.1 donne

$$\sum_{i=1}^{\infty}|\langle x, e_i\rangle|^2 = \|P_F(x)\|^2 \leq \|x\|^2$$

Cela prouve l'assertion (a). Quant à l'assertion (b), elle se démontre en remarquant que pour $y \in E$

$$\langle P_{F_n}(x), P_{F_n}(y)\rangle = \sum_{i=1}^{n}\langle x, e_i\rangle\overline{\langle y, e_i\rangle}$$

et par passage à la limite quand n tend vers l'infini. □

Définition 1.4.4. *On dit qu'une famille $\{ x_i, i \in I \}$ d'éléments d'un espace de Hilbert E, est totale dans E si le sous-espace de Hilbert qu'elle engendre est égal à E.*
Une suite orthonormée et totale dans E est appelée une base hilbertienne de E.

Définition 1.4.5. *Un espace de Hilbert est dit séparable, ou de dimension dénombrable, s'il contient une famille dénombrable totale.*

Un espace de Hilbert qui possède une base hilbertienne est évidemment séparable. Inversement, on a

Théorème 1.4.6. *Si un espace de Hilbert est séparable, alors il possède une base hilbertienne.*

Démonstration. Soit E un espace de Hilbert séparable et (f_n), $n \in \mathbb{N}$, une suite totale. Nous allons construire à partir de cette suite une base hilbertienne.

En enlevant, par récurrence, les f_i qui sont combinaisons linéaires de ceux qui les précèdent, on peut supposer que la suite est formée d'éléments linéairement indépendants.

Pour orthogonaliser cette suite, il suffit alors de procéder à une nouvelle récurrence en définissant une suite (e_n), $n \in \mathbb{N}$, par

$$e_1 = f_1, \quad e_2 = f_2 - P_{F_1}(f_2), \dots, e_n = f_n - P_{F_{n-1}}(f_n)$$

où F_n désigne bien sûr le sous-espace de E engendré par les n premiers vecteurs de la suite (f_n). Par construction même, la suite (e_n), $(n \geq 1)$, est formée de vecteurs deux à deux orthogonaux et, pour tout entier n, le sous-espace vectoriel engendré par $\{ e_1, e_2, \dots, e_n \}$ est égal à F_n. Les éléments (e_n) forment donc une suite totale dans E. En divisant chacun d'eux par sa norme, on obtient une base hilbertienne de E. □

Le théorème suivant caractérise les bases hilbertiennes.

Théorème 1.4.7. *Soient E un espace de Hilbert et (e_n) une suite orthonormée dans E. Les assertions suivantes sont équivalentes :*

(a) *La suite $\{ e_n, n \geq 1 \}$ est une base hilbertienne de E.*

(b) *Tout élément x de E s'écrit sous la forme $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$*

(c) *Quels que soient x et y dans E, on a*

$$\langle x, y \rangle = \sum_{n=1}^{\infty} \langle x, e_n \rangle \overline{\langle y, e_n \rangle}$$

(d) *Pout tout $x \in E$, on a l'égalité de Parseval*

$$\|x\|^2 = \sum_{n \geq 1} |\langle x, e_n \rangle|^2$$

Démonstration. Soit F le sous-espace de Hilbert engendré par la suite (e_n) et soit P_F l'opérateur de projection orthogonale sur F. La suite (e_n) est une base hilbertienne si et seulement si $F = E$, c'est-à-dire $P_F = I$, et le théorème 1.4.3 montre que (a) est équivalent à (b) et implique (c) et (d).

L'assertion (d) traduit le fait que pour tout x dans E, on a l'égalité $\|x\|^2 = \|P_F x\|^2$ et par suite $\|x - P_F(x)\| = 0$. Cela veut dire que $P_F = I$ et donc $F = E$. □

Corollaire 1.4.8. *Tout espace de Hilbert séparable de dimension infinie est isomorphe à* $\ell^2(\mathbb{N})$.

Démonstration. En effet, l'application qui, à un élément x de E, associe la suite de ses composantes $(\langle x, e_i \rangle)$, suivant la base hilbertienne (e_i), est un isomorphisme de E sur $\ell^2(\mathbb{N})$. □

Exemple 1.4.9. - Dans l'espace $\ell^2(\mathbb{N})$, considérons la famille $\{ e_n, n \in \mathbb{N} \}$ où e_n est la suite dont tous les termes sont nuls sauf le $n$ième qui vaut 1. Cela veut dire que e_n représente la suite (δ_{nk}), $k \in \mathbb{N}$, où δ_{nk} est l'indice de Kronecker[13] qui vaut 1 si $k = n$ et 0 sinon. Il suit directement du théorème précédent que cette famille forme une base hilbertienne de l'espace $\ell^2(\mathbb{N})$, dite *la base canonique de* $\ell^2(\mathbb{N})$. De même, dans l'espace de Hilbert $L^2[0, 2\pi]$ muni du produit scalaire

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)\overline{g(t)}\, dt$$

la suite définie par $e_n(t) = e^{int}$, $n \in \mathbb{Z}$, est une base hilbertienne. Nous reviendrons sur cet exemple et sur d'autres au chapitre II.

Remarque 1.4.10. -

• Il ressort du corollaire précédent que l'exemple fondamental d'espace de Hilbert qu'est $\ell^2(\mathbb{N})$, est le modèle des espaces de Hilbert possédant une base dénombrable. Dans de tels espaces, les formules de la géométrie euclidienne, telle que l'écriture d'un vecteur comme la somme de ses composantes dans une base orthonormée, restent valables. Cependant, cela n'a pu se faire qu'en se préoccupant de questions de convergence.

• Soit (e_i) une base hilbertienne de E. Pour tout x dans E, on a

$$\left\| x - \sum_{i=1}^{n} \langle x, e_i \rangle \right\|^2 = \sum_{i=n+1}^{\infty} |\langle x, e_i \rangle|^2$$

Le second membre de cette égalité est le reste d'une série convergente, et donc tend vers zéro quand n tend vers l'infini. Ainsi, $\sum_{i=1}^{n} \langle x, e_i \rangle e_i$ représente une approximation de l'élément x, d'autant plus précise que n est grand.

• A travers la preuve du théorème 4.6, nous reconnaissons dans le procédé qui a permis de construire la suite orthonormée $(e_i)_{i \in \mathbb{N}}$ à partir de la suite $(f_i)_{i \in \mathbb{N}}$, une extension d'un procédé bien connu dans le cadre des

[13] Leopold KRONECKER (1823-1891) est un mathématicien allemand. Il s'est intéressé principalement à la théorie des Equations Algébriques et à la Théorie des Nombres. Il a demandé avec insistance que toutes les mathématiques puissent être fondées sur les nombres entiers. Il disait à ce propos : "Dieu a créé les nombes entiers, tout le reste est l'œuvre de l'Homme".

espaces euclidiens ou hermitiens de dimension finie, appelé *procédé d'orthogonalisation de* Gram[14]- Schmidt[15].

Nous utiliserons ce procédé pour construire quelques exemples (classiques) de bases hilbertiennes et analyser des exemples de problèmes d'approximation qui s'y rattachent.

Somme directe hilbertienne.

Soient F_1 et F_2 deux sous-espaces de Hilbert d'un espace de Hilbert E sur $\mathbb{K}$. Leur somme directe $F_1 \oplus F_2$ est l'ensemble des éléments de la forme : $x = x_1 + x_2$, où x_i appartient à F_i, $1 \leq i \leq 2$. C'est un sous-espace vectoriel. Pour deux élements $x = x_1 + x_2$ et $y = y_1 + y_2$ de $F_1 \oplus F_2$, on vérifie que

$$\langle x, y\rangle = \langle x_1, y_1\rangle + \langle x_2, y_2\rangle$$

On vérifie que cela fait de $F_1 \oplus F_2$ un sous-espace de Hilbert. Il est appelé la somme directe hilbertienne des sous-espaces F_1 et F_2.

On veut généraliser cette construction à une suite de sous-espaces de Hilbert de E, deux à deux orthogonaux.

Proposition 1.4.11. *Soient $\{F_n, n \in \mathbb{N}\}$, une suite de sous-espaces de Hilbert de E et soit F l'ensemble des suites (x_n), où x_n est dans F_n pour tout n, telles que $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$. Pour deux éléments $\mathbf{x} = (x_n)$ et $\mathbf{y} = (y_n)$ dans F, on pose*

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{\infty} \langle x_i, y_i\rangle$$

Alors, $\langle , \rangle$ est un produit scalaire sur F, la norme associée étant

$$\|\mathbf{x}\| = \left[\sum_{i=1}^{\infty} \|x_i\|^2\right]^{\frac{1}{2}}$$

Muni de ce produit scalaire, F est un espace de Hilbert.

Démonstration. On remarque d'abord que, si $\mathbf{x} = (x_n)$ et $\mathbf{y} = (y_n)$ sont deux éléments de F, la série $\sum \langle x_i, y_i\rangle$ est absolument convergente, car

$$\left|\sum \langle x_i, y_i\rangle\right| \leq \sum \|x_i\| \, \|y_i\| \leq \left(\sum \|x_i\|^2\right)^{\frac{1}{2}} \left(\sum \|y_i\|^2\right)^{\frac{1}{2}} < \infty$$

On vérifie ensuite, rapidement, que E est complet. □

[14]Jorgen Pedersen GRAM (1850-1916), mathématicien allemand, a étudié la théorie de la fonction gamma et les systèmes orthogonaux.

[15] Ehrard SCHMIDT (1876-1959) est un mathématicien d'origine allemande, il a développé l'Algèbre linéaire et a participé à la fondation de la théorie des Equations Intégrales.

Définition 1.4.12. *Soit (F_n) une suite de sous-espaces de Hilbert de E. L'espace de Hilbert F, construit ci-dessus, est appelé la somme directe hilbertienne des sous-espaces F_n et on note*

$$F = F_1 \oplus F_2 \oplus \cdots \oplus F_n \oplus \cdots$$

Remarque. Si pour tout entier n on prend $E_n = \mathbb{K}$, l'espace E somme hilbertienne des (E_n) est tout simplement $\ell^2(\mathbb{N})$.

EXERCICES

1. Soit $L^2(0,1)$ muni du produit scalaire $\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\,dx$. Pour tout entier relatif n, on pose $\phi_n(x) = e^{2i\pi nx}$.
(a)Montrer que la suite $(\phi_n)_{n\in\mathbb{Z}}$ est une base hilbertienne de $L^2[0,1]$. Soit F_0 l'ensemble de tous les éléments $f = \sum c_n\phi_n$ de $L^2[0,1]$ tels que $c_{2n+1} = 0$ pour tout $n \in \mathbb{Z}$. Soit F_1 l'ensemble des éléments $g = \sum c_n\phi_n$ de $L^2(0,1)$ tels que $c_{2n} = (1+|n|)c_{2n+1}$ pour tout $n \in \mathbb{Z}$.
(b) Montrer que F_0 et F_1 sont deux sous-espaces fermés de $L^2(0,1)$.
(c) Montrer que $F_0 + F_1$ est dense dans $L^2(0,1)$ et que de plus si $h = \sum c_n\phi_n$ appartient à $F_0 + F_1$ alors

$$\sum_{n\in\mathbb{Z}} n^2|c_{2n+1}|^2 < \infty$$

(d) Conclure que la somme de deux sous-espaces fermés d'un espace de Hilbert n'est pas nécessairement fermée.

2. Soit D le disque unité ouvert de $\mathbb{C}$, soit $\mathcal{O}(D)$ l'ensemble des fonctions holomorphes dans D et soit E l'ensemble des fonctions $f \in \mathcal{O}(D)$ telles que

$$\|f\|^2 = \iint_D |f(z)|^2 dxdy < \infty, \quad \text{où} \quad z = x + iy$$

Pour f et g dans E, on pose : $\langle f, g\rangle = \iint_D f(z)\overline{g(z)}dxdy$.
(a) Montrer que, pour tout $z_0 \in D$ et $f \in E$, on a

$$|f(z_0)| \le \frac{1}{\sqrt{\pi r_0^2}}\|f\|, \quad \text{où} \quad r_0 = \operatorname{dist}(z_0, \partial D)$$

(b) En déduire que E est un espace de Hilbert.
On pose, pour tout entier $n \ge 0$, $e_n(z) = \sqrt{(n+1)/\pi}\,z^n$.
(c) Vérifier que (e_n) est une suite orthonormée dans E. Calculer le produit scalaire $\langle f, e_n\rangle$, pour $f \in E$.
(d) Soit $f \in E$ et $0 < r < 1$ et soit $D(0,r)$ le disque centré à l'origine et de rayon r. En considérant l'intégrale

$$\iint_{D(0,r)} |f(z)|^2 dxdy$$

montrer que

$$\|f\|^2 = \pi \sum_{n\geq 0} \frac{|f^{(n)}(0)|^2}{(n+1)(n!)^2}$$

et en déduire que (e_n) est une base hilbertienne de E.

3. Soit $\mathcal{O}(\mathbb{C})$ l'ensemble des fonctions holomorphes dans $\mathbb{C}$, et soit E l'ensemble des fonctions $f \in \mathcal{O}(\mathbb{C})$ telles que

$$\|f\|^2 = \frac{1}{\pi} \int\!\!\int |f(z)|^2 \exp(-|z|^2) dx dy < \infty$$

Pour f et g dans E, on pose

$$\langle f, g \rangle = \frac{1}{\pi} \int\!\!\int f(z)\overline{g(z)} \exp(-|z|^2) dx dy$$

(a) Montrer que si (a_n) sont les coefficients du développement en série entière d'une fonction $f \in E$ alors, $\|f\|^2 = \sum_{n\geq 0} |a_n|^2(n!)$ et pour tout $z \in \mathbb{C}$, on a $|f(z)| \leq \|f\| \ \exp(|z|^2/2)$.

(b) Soit (a_n) une suite de nombres complexes tels que

$$\sum_{n=0}^{\infty} |a_n|^2(n!) < \infty$$

Montrer que la série $\sum a_n z^n$ définit une fonction f de E.

(c) Montrer que E est un espace de Hilbert et que la suite (ϕ_m) définie par $\phi_m(z) = z^m/\sqrt{m!}$, $m \geq 0$, est une base hilbertienne de E.

(d) Montrer que, pour tout a complexe, l'application $f \mapsto f(a)$ est une forme linéaire continue sur E. Soit e_a l'unique élément de E tel que $f(a) = \langle f, e_a \rangle$. Montrer que la famille (e_a), où a varie dans $\mathbb{C}$, est totale dans E. Déteminer e_a.

4. Soit (ϕ_n) une suite orthonormée d'éléments de $L^2([a,b]; dx)$.

(a) Montrer que la suite (ϕ_n) est totale si, et seulement si, pour tout x dans l'intervalle $[a, b]$

$$\sum_{n\in\mathbb{N}} \left(\int_a^x \phi_n(t) dt \right)^2 = x - a \qquad (*)$$

(b) Montrer que (ϕ_n) est totale si et seulement si

$$\sum_{n\in\mathbb{N}} \int_a^b \left| \int_a^x \phi_n(t) dt \right|^2 dx = \frac{(b-a)^2}{2} \qquad (**)$$

Solution : (a) Si (ϕ_n) est une base hilbertiennes le théorème 1.4.7, appliqué à χ_x la fonction caractéristique de $[0, x]$ $(a \leq x \leq b)$, implique

que $x - a = \|\chi_x\|^2 = \sum_{i=1}^{\infty} |\langle \chi_x, \phi_i \rangle|^2$ ce qui est la relation $(*)$. Inversement, supposons que (ϕ_i) est une suite orthonormée satisfaisant à la relation $(*)$, pour tout $x \in [a, b]$. Le théorème projec :ei assure que la série $\sum_{i=1}^{\infty} \langle \chi_x, \phi_i \rangle \phi_i$ est convergente et a pour somme la fonction χ_x et ce pour tout x dans $[a, b]$. Il en résulte que si une fonction f dans $C[a, b]$ est orthogonale à la suite (ϕ_i), alors elle est orthogonale à χ_x, c'est-à-dire que $F(x) = \langle f, \chi_x \rangle = 0$. Ceci étant pour tout x dans $[a, b]$, on en déduit que F est identiquement nulle sur l'intervalle et il en sera de même de sa dérivée f. La suite (ϕ_n) est donc totale dans le sous-espace $C[a, b]$, qui lui-même est dense dans $L^2[a, b]$. Il en résulte que (ϕ_n) est une base hilbertienne.

D'autre part, par intégration terme à terme, la relation $(*)$ implique la relation $(**)$, pour montrer (b) il suffit donc de prouver que cette dernière implique la relation $(*)$. Or, le théorème projec :ei, appliqué à la fonction χ_x, montre que la série $\sum_i |\langle \chi_x, \phi_i \rangle|^2$ est convergente et que sa somme est inférieure ou égale à $\|\chi_x\|^2 = (x - a)$. La fonction définie sur l'intervalle $[a, b]$ par $A(x) = (x - a) - \sum_i |\langle \chi_x, \phi_i \rangle|^2$ est continue et positive, et la relation $(**)$ se traduit par $\int_a^b A(y)\, dy = 0$. Il en résulte que $A(x) = 0$, $\forall x \in [a, b]$, ce qui est la relation $(*)$. □

Chapitre 2

Exemples de bases hilbertiennes

Dans ce chapitre nous allons passer en revue quelques exemples des plus classiques de bases hilbertiennes. Ce sont des fonctions spéciales qui interviennent dans la résolution d'un grand nombre de problèmes de la physique théorique et mathématique. Elles jouent aussi un rôle important dans la théorie de l'approximation et l'analyse numérique.

2.1 Approximation uniforme

Nous rappelons d'abord le remarquable théorème d'approximation de Weierstrass qui établit la densité de l'espace des polynômes dans l'espace $C[0,1]$ pour la norme uniforme. En d'autres termes, la suite des fonctions

$$1, x, x^2, \dots, x^n, \dots$$

est totale dans $C[0,1]$ muni de la norme uniforme. Soit $\mathcal{P}_n$ l'espace vectoriel des polynômes de degré inférieur ou égal à n sur $\mathbb{R}$.

Théorème 2.1.1. *Soit $[a,b]$ un intervalle fermé borné et f une fonction dans $C[a,b]$. Alors pour tout $\epsilon > 0$, il existe un entier $n \in \mathbb{N}$ et un polynôme $p \in \mathcal{P}_n$ tels que*

$$\|f - p\|_\infty < \epsilon$$

Ce théorème a fasciné les mathématiciens qui lui ont donné plusieurs démonstrations. La démonstration originale de Weierstrass ne donne pas un procédé d'approximation pratique et ne fournit pas une majoration commode de la quantité $\|f-p\|_\infty$. On doit à F. Bernstein un procédé explicite d'approximation uniforme d'une fonction f continue sur $[0,1]$ par des polynômes appelés polynômes de Bernstein[1].

[1]Félix Bernstein (1878–1956) fut élève de Cantor, puis de Hilbert et Klein. Il devint professeur de statistique mathématique en 1911 à l'Université de Göttingen. A partir de 1934, il immigra aux Etats Unis d'Amérique où il enseigna dans plusieurs Universités.

Définition 2.1.2. *On appelle polynôme de Bernstein d'ordre n d'une fonction f de* $C[0,1]$ *et on note* $B_n f$*, le polynôme de degré inférieur ou égal à n défini par*

$$B_n f(x) = \sum_{i=0}^{n} f(i/n) \binom{n}{i} x^i (1-x)^{n-i}, \quad n = 1, 2, \ldots$$

Démonstration. Quitte à faire le changement $t = (x-a)/(b-a)$, on peut supposer dans la suite que $a = 0$ et $b = 1$. Désignons par q_{ni} le polynôme défini par

$$q_{ni}(x) = \binom{n}{i} x^i (1-x)^{n-i}, \quad \text{où} \quad i \leq n$$

Les polynômes q_{ni} vérifient les relations

$$\sum_{i=0}^{n} q_{ni}(x) = 1 \qquad (1) \qquad (1)$$

$$\sum_{i=0}^{n} i q_{ni}(x) = nx \qquad (2)$$

$$\sum_{i=0}^{n} i(i-1) q_{ni}(x) = n(n-1)x^2 \qquad (3)$$

Pour le voir on considère la formule de Newton[2]

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}$$

En prenant $y = 1 - x$, on obtient (1). Si l'on dérive la formule de Newton par rapport à x et on multiplie par x, on obtient

$$nx(x+y)^{n-1} = \sum_{i=0}^{n} i \binom{n}{i} x^i y^{n-i}$$

et la relation (2) s'en déduit en prenant $y = 1 - x$. En dérivant deux fois la formule de Newton par rapport à x et en multipliant par x^2, on obtient la relation

$$n(n-1)x^2(x+y)^{n-2} = \sum_{i=0}^{n} i(i-1) \binom{n}{i} x^i y^{n-i}$$

[2]La contribution de Isaac Newton (1642–1727) en mathématiques et en physique est très importante. Il généralisa la formule du binôme, inventa le calcul des fluxions ("Calcul Différentiel") et proposa le principe de la gravitation universelle. Voilà ce qu'il écrivait à la fin de sa vie : "J'ignore sous quel aspect je puis apparaître au monde ; mais, à moi-même, je me fais l'effet de n'avoir été autre chose qu'un garçon jouant sur le rivage, et m'accusant de temps à autre à trouver un cailloux plus poli ou un coquillage plus joli qu'à l'ordinaire, tandis que le grand océan de la Vérité se déroulait devant moi sans que je ne le connusse."

qui implique (3) lorsque l'on prend $y = 1 - x$. Les formules (1), (2) et (3), ainsi prouvées, impliquent la suivante

$$\sum_{i=n}^{n} (i - nx)^2 q_{ni}(x) = nx(1-x) \tag{4}$$

La relation (1) permet d'écrire

$$f(x) - B_n f(x) = \sum_{i=0}^{n} \big[f(x) - f(i/n)\big] q_{ni}(x)$$

et il vient donc, pour tout $x \in [0,1]$,

$$|f(x) - B_n f(x)| \leq \sum_{i=0}^{n} \big|f(x) - f(i/n)\big| q_{ni}(x) \tag{5}$$

Ceci étant, la fonction f est uniformément continue sur $[0,1]$, donc pour tout $\epsilon > 0$ il existe un réel $\delta > 0$ tel que $|f(x) - f(i/n)| < \epsilon/2$ et ce pour tout x tel que $|x - (i/n)| < \delta$. Considérons alors pour tout $x \in [0,1]$ les deux ensembles

$$N' = \{i \in \mathbb{N} : \mid x - (i/n) \mid < \delta\}$$
$$N'' = \{i \in \mathbb{N} : \mid x - (i/n) \mid \geq \delta\}$$

On a

$$\sum_{i \in N'} \big|f(x) - f(i/n)\big| q_{ni}(x) \leq \frac{\epsilon}{2} \sum_{i \in N'} q_{in}(x) \leq \frac{\epsilon}{2} \sum_{i=0}^{n} q_{in}(x) = \frac{\epsilon}{2}$$

En désignant par M le maximum de f sur l'intervalle $[0,1]$ et sachant que pour n dans N'', $(nx - i)^2 \geq n^2\delta^2$, on a aussi d'après (4)

$$\sum_{i \in N''} \big|f(x) - f(i/n)\big| q_{ni}(x) \leq 2M \sum_{i \in N''} q_{ni}(x)$$
$$\leq \frac{2M}{n^2\delta^2} \sum_{i=0}^{n} q_{ni}(x)(i - nx)^2 \leq \frac{2M}{n\delta^2} x(1-x) \leq \frac{M}{2n\delta^2}$$

Il existe un entier $N(\epsilon)$ tel que le dernier terme de l'inégalité ci-dessus soit inférieure à $\epsilon/2$ dès que $n \geq N(\epsilon)$. Donc, pour tout $\epsilon > 0$, et pour n plus grand que $N(\epsilon)$,

$$\mid f(x) - B_n f(x) \mid < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

pour tout $x \in [0,1]$, ce qui établit la convergence uniforme sur $[0,1]$ de la suite $(B_n f)$ vers f. □

Remarques

• Si l'on examine de près la démonstration précédente, on s'aperçoit que l'estimation de la somme (5) dépend essentiellement de l'estimation de (5) pour les fonctions particulières $f_0(x) = 1$, $f_1(x) = x$ et $f_2(x) = x^2$. Cela suggère que la convergence uniforme de la suite des polynômes de Bernstein d'une fonction continue sur $[a, b]$ dépend de la manière dont se comportent les polynômes de Bernstein pour les trois fonctions f_0, f_1 et f_2. Cette constatation, jointe au fait que l'opérateur qui à une fonction continue f associe son polynôme de Bernstein $B_n f$ est positif, est à l'origine des généralisations du théorème par P. P. Korovkin.

• Les polynômes de Bernstein $B_n f$ ne constituent pas une bonne approximation uniforme de f pour n petit. Leur succès vient du fait que, pour n petit, ils conservent les propriétés géométriques globales de f (monotonie, convexité). La figure 2.1, qui représente les graphes de $f(x) = x^3$ et des polynômes de Bernstein $B_2 f$ et $B_5 f$, montre que la converge de $(B_n f)$ est lente .

• Notons aussi que la méthode d'approximation la plus naturelle, à savoir l'interpolation polynomiale, ne donne pas nécessairement de suite de polynômes convergente vers f. Par exemple, Bernstein a montré que si les points d'interpolation sont équidistants dans l'intervalle $[-1, 1]$, avec $x_0 = -1$ et $x_n = 1$, alors pour $f(x) = |x|$, les polynômes d'interpolation de Lagrange divergent en chaque point de $[-1, 1]$ excepté en $x = 0$, $x = -1$ et $x = 1$! Cette situation est connue sous le nom de phénomène de Runge[3].

• Rappelons enfin, que si (p_n) est une suite de polynômes qui converge uniformément sur un intervalle non borné vers une fonction f, alors f est un polynôme. Ainsi, l'idée d'approcher uniformément une fonction continue par des polynômes sur un intervalle non borné est dépourvue d'intérêt. **Le cas des fonctions périodiques**

On appelle polynôme trigonométrique de degré inférieur ou égal à n en la variable réelle x, toute combinaison linéaire (à coefficients complexes) de la forme

$$P(x) = \sum_{m=-n}^{n} c_m e^{imx}, \qquad \text{avec} \ \ c_m \in \mathbb{C}$$

En décomposant e^{imx} en partie réelle et partie imaginaire, il revient au même de dire qu'un polynôme trigonométrique de degré inférieur ou égal

[3]Carl David Tolmé Runge (1856–1927) était le premier à occuper la chaire de mathématiques appliquées, crée à l'Université de Göttingen sous l'impulsion de Félix Klein. Il a été influencé surtout par Weierstrass. Ses travaux ont porté sur la géométrie différentielle, l'algèbre, la théorie des fonctions et surtout les méthodes numériques. Il est considéré parmi ceux qui ont construit un pont entre les mathématiques et les Sciences Techniques.

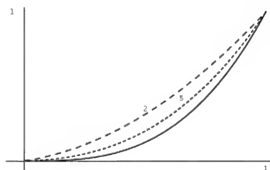


FIG. 2.1 –

à n est une fonction de la forme

$$P(x) = \sum_{m=0}^{n} (a_m \cos mx + b_m \sin mx)$$

Un polynôme trigonométrique est évidemment une fonction continue et périodique de période 2π. En outre, les coefficients c_m sont bien déterminés par les valeurs de P, en effet les relations d'orthogonalité

$$\frac{1}{2\pi} \int_0^{2\pi} e^{imx} e^{-ikx}\, dx = \delta_{km}$$

donnent les formules

$$c_m = \frac{1}{2\pi} \int_0^{2\pi} P(x) e^{-imx}\, dx, \quad -n \leq m \leq n$$

Plus généralement, soit f une fonction à valeurs complexes, 2π-périodique et continue dans l'intervalle $[0, 2\pi]$ (et par suite, en vertu de la périodicité, continue dans $\mathbb{R}$). On appelle coefficients de Fourier de f les nombres complexes

$$c_m(f) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-imx}\, dx, \quad m \in \mathbb{Z}$$

On est naturellement amené, par analogie avec ce qui a été vu plus haut, à associer à f les polynômes trigonométriques

$$S_nf(x)=\sum_{m=-n}^{n}c_m(f)e^{imx},\quad n\in\mathbb{N}$$

où les $c_m(f)$ sont les coefficients de Fourier de f. On dit que S_nf est le polynôme de Fourier de f ; il est de degré n si l'un au moins des coefficients c_n ou c_{-n} est non nul. C'est aussi la somme partielle de la série

$$\sum_{m\in\mathbb{Z}}c_m(f)e^{imx}$$

dite série de Fourier de f. On peut voir à l'exercice 1 que S_nf s'écrit à l'aide du noyau de Dirichlet[4] D_n, sous la forme

$$S_nf=\frac{1}{2\pi}\int_0^{2\pi}f(t)D_n(x-t)\,dt,\quad \text{où}\quad D_n(t)=\frac{\sin((2n+1)\frac{t}{2})}{\sin(\frac{t}{2})}$$

On pourrait espérer que S_nf converge uniformément sur $[0,2\pi]$ (ou au moins simplement) vers f, mais on connaît des exemples de fonctions périodiques continues, dont la série de Fourier est divergente en certains points. Cela provient du fait que

$$\sup_n\int_0^{2\pi}|D_n(t)|\,dt=+\infty$$

Ce problème peut être surmonté en considérant les sommes de Césaro, (E. Césaro : 1859-1906). La somme de Césaro d'ordre $n\geq 1$ est par définition la moyenne arithmétique des n premiers termes de la suite (S_nf), C'est-à-dire

$$\sigma_n(f):=\frac{S_0f+S_1f+\cdots+S_{n-1}f}{n}$$

Il est plus commode d'écrire $\sigma_n(f)$ sous forme intégrale. La relation

$$\sin(j+\frac{1}{2})u-\sin(\frac{u}{2})=\frac{1}{2}[\cos ju-\cos(j+1)u]$$

permet d'écrire

$$\begin{aligned}\sum_{j=0}^{n-1}\sin(j+\frac{1}{2})u\sin(\frac{u}{2})&=\frac{1}{2}\sum_{j=0}^{n-1}[\cos ju-\cos(j+1)u]\\&=\frac{1}{2}(1-\cos nu)=\sin^2(\frac{nu}{2})\end{aligned}$$

[4]Peter-Gustave Lejeune Dirichlet (1805–1859) forme, avec ses amis Jacobi et Kummer, la première génération des mathématiciens allemands après Gauss. Le sujet de prédilection de Dirichlet, pendant toute sa carrière, a été la théorie des nombres. On lui doit deux outils puissants : le célèbre "principe des tiroirs" et les "séries de Dirichlet".

et on en déduit alors

$$\begin{aligned}\sigma_n(f)(x) &= \frac{1}{2\pi n}\int_0^{2\pi} f(t)\left[\sum_{j=0}^{n-1}\frac{\sin((2j+1)\frac{x-t}{2})}{\sin(\frac{x-t}{2})}\right]dt\\ &= \frac{1}{2\pi n}\int_0^{2\pi} f(t)\frac{\sin^2(n(x-t)/2)}{\sin^2(\frac{x-t}{2})}\,dt\end{aligned}$$

Ainsi, $\sigma_n(f)$ apparaît comme le produit de convolution de la fonction f avec la fonction

$$F_n(x) = \frac{1}{n}\frac{\sin^2(\frac{nx}{2})}{\sin^2(\frac{x}{2})}$$

appelée parfois noyau de Féjer, du nom du mathématicien L. Féjer (1880–1959). En vertu de la périodicité, on peut écrire aussi

$$\sigma_n(f)(x) = \frac{1}{2\pi n}\int_{-\pi}^{\pi} f(x-t)\frac{\sin^2(\frac{nt}{2})}{\sin^2(\frac{t}{2})}\,dt$$

Contrairement au noyau de Dirichlet, le noyau de Féjer F_n est positif, il possède de plus les deux propriétés suivantes (voir exercice 2)

$$\frac{1}{2\pi}\int_0^{2\pi} F_n(x)\,dx = 1, \quad \text{et pour } \delta > 0,\ \lim_{n\to 0}\int_\delta^\pi F_n(x)\,dx = 0$$

Théorème 2.1.3. *Toute fonction continue, 2π-périodique, est limite uniforme sur $[0, 2\pi]$ d'une suite de polynômes trigonométriques.*

Démonstration. Nous allons montrer que la suite $(\sigma_n(f))$ des sommes de Césaro converge uniformément sur $[0, 2\pi]$ vers f. En effet, compte tenu des propriétés du noyau de Féjer, on a

$$\sigma_n(f)(x) - f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}[f(x-t) - f(x)]F_n(t)\,dt$$

Puisque f est uniformément continue sur $[-\pi, \pi]$, pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que pour $|t| < \delta$, on ait

$$\sup_{-\pi\le x\le\pi} \mid f(x-t) - f(x) \mid < \epsilon$$

le nombre δ étant ainsi choisi, il existe un entier n_0 tel que

$$\int_{-\pi}^{-\delta} F_n(t)\,dt + \int_\delta^\pi F_n(t)\,dt < \frac{\epsilon}{2\|f\|_\infty}, \quad \text{pour } n \ge n_0$$

Il en résulte alors que, pour tout $n \ge n_0$ et tout $x \in [0, 2\pi]$,

$$|f(x) - \sigma_n(f)(x)| \le \epsilon\int_{-\delta}^{\delta} F_n(t)\,dt + \epsilon \le 2\epsilon$$

La preuve du théorème est ainsi terminée. □

On s'intéresse maintenant à la question de séparabilité d'un espace de Hilbert. Plus précisément, considérons un espace de Hilbert de la forme $E = L^2(I, \mu)$ où I est un intervalle de $\mathbb{R}$ et μ une mesure absolument continue par rapport à la mesure de Lebesgue[5] et dont la densité ω est supposée continue et strictement positive sur I. On supposera que les intégrales $\int_I x^{2n}\omega(x)dx$ sont finies pour tout entier $n \in \mathbb{N}$. Cette hypothèse assure que les fonctions polynômes sont dans l'espace $L^2(I, \mu)$.

Théorème 2.1.4. *Si l'intervalle I est borné, alors l'espace $L^2(I, \mu)$ est séparable.*

Démonstration. En effet, le théorème de Weierstrass dit que toute fonction continue sur l'adhérence de I est limite uniforme de polynômes ; ensuite la convergence uniforme entraîne la convergence au sens $L^2(I, \mu)$ car pour toute fonction continue sur $\overline{I}$ on a

$$\|f\|^2_{L^2(I,\mu)} = \int_I |f(x)|^2 \omega(x)dx \leq \|f\|^2_\infty \int_I \omega(x)dx$$

enfin, les fonctions continues sont partout denses dans $L^2(I, \mu)$. □

Dans ce cas, on peut appliquer le procédé d'orthogonalisation de Gram-Schmidt à la suite (f_n) définie par $f_n(x) = x^n$ et il est clair que les éléments (p_n) ainsi construits sont des polynômes de degré n qui constituent une base orthogonale de $L^2(I, \mu)$. Dans le cas général (I non borné), on a le résultat suivant

Théorème 2.1.5. *Supposons qu'il existe un réel $r > 0$ tel que*

$$\int_I e^{r|x|}\omega(x)dx < \infty$$

Alors, la suite (f_n), définie par $f_n(x) = x^n$, $n \in \mathbb{N}$, est totale dans $L^2(I, \mu)$.

Démonstration. Soit f une fonction de $L^2(I, \mu)$ telle que, pour tout $n \in \mathbb{N}$, on ait

$$\int_I f(x)x^n\omega(x)dx = 0$$

Il s'agit de vérifier que $f = 0$. Posons

$$F(z) = \int_I f(x)e^{-izx}\omega(x)dx$$

[5]Henri LEBESGUE (1875-1941) a été l'auteur de nombreux travaux d'Analyse des fonctions d'une variable réelle. Sa Théorie de l'Intégration reste aujourd'hui la référence en ce domaine.

La fonction F est bien définie sur $\mathcal{O}_r = \{z \in \mathbb{C} \mid |\Im m(z)| < r/2\}$. Le théorème de convergence dominée montre qu'elle est holomorphe sur l'ouvert $\mathcal{O}_r$ et pour tout $n \in \mathbb{N}$,

$$F^{(n)}(z) = \int_I f(x)(-ix)^n e^{(-izx)}\omega(x)dx$$

On en déduit que toutes les dérivées de F sont nulles en 0, ce qui implique que F est identiquement nulle. Mais F n'est autre que la transformée de Fourier de la fonction $f.\omega$ prolongée à $\mathbb{R}$ par 0 en dehors de I ; on en déduit que f est identiquement nulle. □

Conséquence. Soit ω une fonction vérifiant l'hypothèse du théorème 2.1.5. Le procédé d'orthogonalisation de Gram-Schmidt, appliqué à la suite définie par $f_n(x) = x^n$, donne naissance à une suite de polynômes (p_n), $\deg p_n = n$, qui constituent une base orthogonale de l'espace $L^2(I, \omega(x)dx)$. En fait, cette base est déterminée de façon unique (à une constante multiplicative près) par le poids ω, plus précisémment on a

Théorème 2.1.6. *Soit (q_n) une suite de polynômes deux à deux orthogonaux dans $L^2(I, \omega(x)dx)$ et tels que $\deg q_n = n$, pour tout n. Alors*

(a) *Tout polynôme de degré n peut être représenté sous la forme d'une combinaison linéaire des polynômes q_m, avec $0 \leq m \leq n$, et on a*

(b) $\displaystyle\int_I q_n(x)x^m\,\omega(x)dx = 0, \quad \text{pour} \quad m < n.$

(c) *Si p_n est une autre suite de polynômes, deux à deux orthogonaux dans $L^2(I, \omega(x)dx)$, telle que $\deg p_n = n$ pour tout n, il existe une constante c_n telle que $p_n = c_n q_n$.*

Démonstration. La propriété (a) est en fait vraie même si les polynômes q_n ne sont pas deux à deux orthogonaux. En effet, il est clair qu'il existe des constantes c_0 et c_1 telles que

$$x = c_0 q_0(x) + c_1 q_1(x)$$

Supposons par récurrence que tout polynôme de degré inférieur ou égal à m peut être représenté par une combinaison linéaire des polynômes (q_j), avec $0 \leq j \leq m$. Le polynôme q_{m+1}, étant de degré $m+1$, s'écrit sous la forme

$$q_{m+1}(x) = c_{m+1}x^{m+1} + Q_m(x), \quad \text{avec } c_{m+1} \neq 0 \text{ et } \deg Q_m \leq m$$

l'hypothèse de récurrence montre que Q_m est combinaison linéaire des polynômes q_j, $0 \leq j \leq m$, et il en résulte que le monôme

$$c_{m+1}x^{m+1} = [q_{m+1}(x) - Q_m(x)]$$

est une combinaison linéaire des polynômes q_j, avec $0 \leq j \leq m+1$. Ainsi l'assertion (a) est vraie pour les monômes x^m et on en déduit facilement qu'elle est vraie pour tout polynôme.

Montrons la relation (b). D'après (a), pour tout $m < n$ on peut écrire le monôme x^m sous la forme

$$x^m = c_0 q_0(x) + c_1 q_1(x) + \cdots + c_m q_m(x)$$

En multipliant les deux membres de cette égalité par q_n et en intégrant, on en déduit immédiatement la propriété (b) et ce grâce à l'orthogonalité des polynômes (q_j).

Démontrons l'assertion (c). Le polynôme p_n est de degré n, donc d'après la propriété (a), il peut s'écrire sous la forme

$$p_n(x) = c_0 q_0(x) + c_1 q_1(x) + \cdots + c_n q_n(x)$$

mais p_n est orthogonal à tout polynôme de degré inférieur ou égal à $n-1$, en particulier, pour tout $j \leq n-1$

$$0 = \int_I p_n(x) q_j(x) \omega(x)\, dx = c_j \int_I q_j^2(x) \omega(x) dx$$

Il en résulte que $p_n = c_n q_n$. □

Remarque 2.1.7. - La relation (c) traduit le fait que les polynômes (q_n), obtenus par le procédé d'orthogonalisation de Gram-Schmidt appliqué à la suite des monômes x^n, $n \in \mathbb{N}$, sont définis (à une constante multiplicative près) de façon unique par le poids ω. On en déduit notamment la propriété de parité suivante

Corollaire 2.1.8. *Si l'intervalle I est symétrique par rapport à l'origine et si la fonction ω est paire, alors les polynômes q_n ont la parité de n, c'est-à-dire $q_n(-x) = (-1)^n q_n(x)$, pour tout n.*

Démonstration. Il suffit de remarquer que pour tout n, la fonction $\tilde{q}_n$ définie par $\tilde{q}_n(x) = q_n(-x)$ est un polynôme de degré n et que, pour $n \neq m$

$$\int_I \tilde{q}_n(x) \tilde{q}_m(x) \omega(x)\, dx = \int_I q_n(x) q_m(x) \omega(x) dx = 0$$

la relation (c) du théorème précédent montre alors qu'il existe une constante c_n telle que $\tilde{q}_n = c_n q_n$. En examinant les termes de plus haut degré on trouve que $c_n = (-1)^n$. □

Les paragraphes qui suivent sont consacrés à la présentation de quelques exemples classiques de bases hilbertiennes.

EXERCICES

1. On désigne par D_p le noyau de Dirichlet : $D_p(x) = \sum_{n=-p}^{p} e^{inx}$.
(a) Montrer que D_p est 2π-périodique, paire dont l'intégrale sur $[0, 2\pi]$ vaut 2π.
(b) Montrer que D_p peut s'écrire sous la forme

$$D_p(x) = \frac{\sin(p + \frac{1}{2})x}{\sin(\frac{x}{2})}$$

(c) Soit f une fonction continue et 2π-périodique. Montrer que son polynôme de Fourier de degré p, $S_p f = \sum_{-p}^{p} c_n(f) e^{inx}$, s'écrit

$$S_p f(x) = \frac{1}{2\pi} \int_0^{2\pi} f(t) D_p(x - t)\, dt$$
$$S_p f(x) = \frac{1}{2\pi} \int_0^{2\pi} f(x - t) D_p(t)\, dt$$

(d) En appliquant l'inégalité $\sin(t/2) \leq t/2$, $(0 \leq t \leq \pi)$, montrer que

$$\int_{-\pi}^{\pi} |D_p(t)|\, dt \geq \frac{2}{\pi} \int_0^{\pi} \frac{|\sin(p + \frac{1}{2})t|}{t}\, dt$$
$$\geq \frac{2}{\pi} \int_0^{(p+\frac{1}{2})\pi} \frac{|\sin(t)|}{t}\, dt$$

et en déduire que cette dernière quantité croît indéfiniment lorsque l'entier n tend vers l'infini.

2. Montrer que le noyau de Féjer F_n est égal à la moyenne arithmétique des n premiers noyaux de Dirichlet

$$F_n(x) = \frac{D_0(x) + D_1(x) + \cdots + D_{n-1}(x)}{n}$$

En déduire que : $\int_0^{2\pi} F_n(x)\, dx = 2\pi$.
Soit $\delta > 0$, montrer que

$$\int_{\delta}^{\pi} F_n(x)\, dx \leq \frac{1}{n \sin^2(\frac{\delta}{2})}$$

et en déduire que

$$\lim_{n \to \infty} \int_{\delta}^{\pi} F_n(x)\, dx = 0$$

3. En utilisant le procédé d'orthogonalisation de Gram-Schmidt, calculer les quatre premiers éléments de la base orthogonale dans chacun des cas suivants
(a) L'ensemble $\{x^n, n \geq 0\}$ dans l'espace $L^2((-1,1), \sqrt{|x|}dx)$
(b) L'ensemble $\{e^{-nx}, n \geq 1\}$ dans l'espace $L^2((0,\infty), dx)$
(c) L'ensemble $\{x^n, n \geq 0\}$ dans l'espace $L^2(\mathbb{R}, e^{-|x|}dx)$.

4. Soit ω une fonction continue sur l'intervalle $[a,b]$ et strictement positive. Soit (p_n) une suite de polynômes orthogonaux sur l'intervalle $[a,b]$ relativement à la mesure de densité ω par rapport à la mesure de Lebesgue

$$\int_a^b p_n(x)p_m(x)\omega(x)\,dx = 0, \quad \forall\, m \neq n$$

(a)Montrer que les racines de p_n sont simples.
(b)Montrer que si $x \notin [a,b]$, on a

$$\int_a^b \frac{p_n^2(s)}{s-x}\omega(s)\,ds = p_n(x)\int_a^b \frac{p_n(s)}{s-x}\omega(s)\,ds \qquad (*)$$

(c)En déduire que toutes les racines de p_n sont dans $[a,b]$.
Solution. (a) Soient $x_1, x_2, \ldots, x_m$ les racines de p qui sont de multiplicité impaire et posons $P(x) = \prod_{i=1}^m (x - x_i)$. Alors,

$$\int_a^b p_n(x)P(x)\omega(x)\,dx > 0$$

Mais si $m < n$, P et p_n seraient orthogonaux et le premier membre de l'inégalité ci-dessus serait alors nul. cela est impossible et donc $m = n$, c'est-à-dire que les racines de p_n sont toutes simples.
(b) Soit x en dehors de l'intervalle $[a,b]$. Le premier membre de $(*)$ peut sécrire sous la forme

$$\int_a^b p_n(s)\frac{p_n(s)-p_n(x)}{s-x}\omega(s)\,ds + p_n(x)\int_a^b \frac{p_n(s)}{s-x}\omega(s)\,ds$$

L'expression $[p_n(s) - p_n(x)/(s-x)]$ est un polynôme en s de degré $(n-1)$, ce polynôme est donc orthogonal à p_n. On a donc obtenu la formule annoncée.
(c) Si x est un réel en dehors de $[a,b]$, le premier membre de $(*)$ garde un signe constant et par suite $p_n(x) \neq 0$. Si $x = \sigma + i\tau$, avec $\tau \neq 0$, la partie imaginaire du premier membre de $(*)$ s'écrit

$$\tau\int_a^b \frac{p_n^2(s)\omega(s)}{(s-\sigma)^2+\tau^2}\,ds$$

et est donc non nulle. Il s'ensuit, dans ce cas aussi, que $p_n(x) \neq 0$. □

5. Considérons dans l'espace de Hilbert $L^2((a,b); w(x)dx)$, une base orthogonale formée de polynômes (p_n). Montrer que le poids ω est déterminé de façon unique par la suite (p_n).

2.2 Séries de Fourier

Notons E l'espace vectoriel $L^2[0,2\pi]$ des (classes de) fonctions de carré sommables sur $[0,2\pi]$ pour la mesure de Lebesgue. Une classe de fonctions qui appartient à E est représentée par une fonction $f : [0,2\pi] \to \mathbb{C}$ mesurable et de carré intégrable sur $[0,2\pi]$. Il sera commode de prolonger f en une fonction $\tilde{f}$, 2π-périodique et mesurable ; il suffit de modifier éventuellement f en un point pour que $f(0) = f(2\pi)$, et de poser $\tilde{f}(x + 2k\pi) = f(x)$ pour $x \in [0,2\pi]$ et $k \in \mathbb{Z}$. Inversement, une fonction $g : \mathbb{R} \to \mathbb{C}$, 2π-périodique mesurable et de carré intégrable sur $[0,2\pi]$ définit, par restriction, un élément de E. Sur E le produit scalaire et la norme sont donnés par

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt$$

$$\|f\|_2^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\,dt, \quad \text{pour} \quad f, g \in E$$

Pour tout $n \in \mathbb{Z}$, on désigne par e_n la fonction définie sur $\mathbb{R}$ à valeurs dans $\mathbb{C}$ par $e_n(x) = e^{inx}$. Les fonctions (e_n) constituent une famille orthonormale et nous allons montrer que c'est une base hilbertienne de $L^2[0,2\pi]$.

En effet, on sait que les fonctions 2π-périodiques et continues sur $[0,2\pi]$ constituent un sous-espace vectoriel dense de $L^2[0,2\pi]$, cela veut dire que pour tout élément f de $L^2[0,2\pi]$ et tout $\epsilon > 0$, il existe une fonction g, 2π-périodique et continue sur $[0,2\pi]$ telle que $\|f - g\|_2 \leq \epsilon$. Le théorème 1.3 montre qu'il existe un polynôme trigonométrique P tel que $\|g - P\|_\infty \leq \epsilon$. A fortiori, on a $\|g - P\|_2 \leq \epsilon$ (c'est ce qu'on exprime en disant que la convergence uniforme implique la convergence en moyenne quadratique). On en déduit, en utilisant l'inégalité triangulaire, que

$$\|f - P\|_2 \leq 2\epsilon$$

Les polynômes trigonométriques sont donc denses dans l'espace $L^2[0,2\pi]$, ce qui traduit le fait que la suite (e_n), $n \in \mathbb{Z}$, est totale dans $L^2[0,2\pi]$ et en est donc une base hilbertienne. En vertu du théorème 1.4.7, chapitre I, on peut énoncer

Théorème 2.2.1. *La suite (e_n), $n \in \mathbb{Z}$, est une base hilbertienne de l'espace $L^2[0,2\pi]$. La série de Fourier d'une fonction f de $L^2[0,2\pi]$ est convergente*

dans $L^2[0,2\pi]$ et sa somme est égale à f, c'est-à-dire

$$\lim_{n\to+\infty}\Big\|f-\sum_{|k|\leq n}c_k(f)e_k\Big\|=0 \quad \text{où} \quad c_n(f)=\langle f,e_n\rangle$$

De plus, pour deux fonctions f et g dans $L^2[0,2\pi]$, on a les formules de Parseval

$$\|f\|_2^2=\sum_{n\in\mathbb{Z}}|c_n(f)|^2 \quad \text{et} \quad \frac{1}{2\pi}\int_0^{2\pi}f(t)\overline{g(t)}\,dt=\sum_{n\in\mathbb{Z}}c_n(f)\overline{c_n(g)}$$

Lorsque l'on s'occupe de fonctions réelles et que l'on utilise les coefficients de Fourier $a_n(f)$ et $b_n(f)$, où l'on a posé

$$a_0(f)=\frac{1}{2\pi}\int_0^{2\pi}f(t)\,dt,$$

et pour $n>0$

$$a_n(f)=\frac{1}{\pi}\int_0^{2\pi}f(t)\cos nt\,dt \quad \text{et} \quad b_n(f)=\frac{1}{\pi}\int_0^{2\pi}f(t)\sin nt\,dt$$

les relations précédentes s'écrivent :

$$\|f\|_2^2=|a_0(f)|^2+\frac{1}{2}\sum_{n\geq1}\left[|a_n(f)|^2+|b_n(f)|^2\right]$$

$$\langle f,g\rangle=a_0(f)\overline{a_0(g)}+\frac{1}{2}\sum_{n\geq1}\left[a_n(f)\overline{a_n(g)}+b_n(f)\overline{b_n(g)}\right]$$

Une conséquence de ce théorème est donnée par le corollaire suivant

Corollaire 2.2.2. *Soit f un élément de $L^2[0,2\pi]$. Dans l'ensemble des polynô-mes trigonométriques P de degré inférieur ou égal à n, le polynôme de Fourier $S_nf=\sum_{|k|\leq n}c_k(f)e_k$ réalise le minimum de $\|f-P\|_2$, et il est le seul à avoir cette propriété.*

Démonstration. L'ensemble $\mathcal{P}_n$ des polynômes trigonométriques de degré inférieur ou égal à n est un sous-espace vectoriel de $L^2[0,2\pi]$ engendré par la famille $\{e_k,|k|\leq n\}$. Le polynôme de Fourier S_nf est la projection orthogonale de f sur $\mathcal{P}_n$, d'où le résultat. □

Remarque 2.2.3. - En ce qui concerne la convergence des séries de Fourier, comme séries de fonctions, on peut citer le résultat de Carlson (1965) qui dit que la série de Fourier d'une fonction $f\in L^2[0,2\pi]$ converge et a pour somme $f(x)$ pour presque tout $x\in[0,2\pi]$. Ce résultat est valable pour une fonction f dans $L^p[0,2\pi]$, pour $1<p<\infty$ (R. A. Hunt, 1967). On sait par ailleurs qu'il existe des fonctions f intégrables sur $[0,2\pi]$ dont la série de Fourier diverge en tout point (Kolmogorov, 1925).

Exemple 2.2.4. - Considérons la fonction f définie par

$$f(x) = \begin{cases} -1, & \text{pour } -\pi \leq x < 0\,; \\ 0, & \text{pour } x = 0\,; \\ 1, & \text{pour } 0 < x < \pi. \end{cases}$$

et par $f(x + 2\pi) = f(x)$. C'est une fonction périodique de période 2π et impaire, donc $a_n(f) = 0$ pour $n \geq 0$ et un calcul simple donne

$$b_n(f) = \frac{2}{\pi}\int_0^{\pi} \sin(nx)\,dx = \begin{cases} 4/(\pi n), & \text{pour } n \text{ impair}\,; \\ 0, & \text{pour } n \text{ pair}. \end{cases}$$

Le polynôme de Fourier de degré $2n - 1$ de f est donné par

$$S_{2n-1}f(x) = \frac{4}{\pi}\sum_{m=1}^{n} \frac{\sin(2m-1)x}{2m-1}$$

Il représente la meilleure approximation en moyenne quadratique de f par des polynômes trigonométriques de degré inférieur ou égal à $2n - 1$. La formule de Parseval donne dans ce cas

$$\frac{\pi^2}{8} = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}$$

Sur la figure 2.2 on a représenté les graphes de f et des polynômes $S_n f$, pour $n = 1, 2$ et 3.

Exemple 2.2.5. - Soit f la fonction 2π-périodique, paire et égale à $\pi - x$ dans l'intervalle $[0, \pi]$. On vérifie que f est continue sur $\mathbb{R}$ et la parité fait que $b_m(f) = 0$, pour tout $m \geq 1$. Justement, en tenant compte de la parité et après une intégration par parties, on trouve que $a_0(f) = \pi/2$ et pour $m \geq 1$

$$\begin{aligned} a_m(f) &= \frac{2}{\pi}\int_0^{\pi} (\pi - x)\cos mx\,dx \\ &= \frac{2}{\pi m^2}(1 - \cos m\pi) = \begin{cases} 0, & \text{si } m \text{ est pair}\,; \\ 4/(\pi m^2), & \text{si } m \text{ est impair}. \end{cases} \end{aligned}$$

On a donc

$$f(x) = \frac{\pi}{2} + \frac{4}{\pi}\sum_{1}^{\infty} \frac{\cos(2m-1)x}{(2m-1)^2}$$

où la convergence est, cette fois-ci, non seulement en moyenne quadratique, mais aussi uniforme sur $\mathbb{R}$. La formule de Parseval donne

$$\frac{\pi^4}{96} = \sum_{m=1}^{\infty} \frac{1}{(2m-1)^4}$$

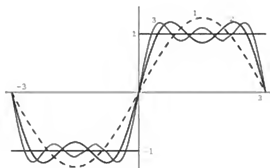


FIG. 2.2 —

Sur la figure 2.3 on a représenté les graphes de f et de son polynôme de Fourier d'ordre 5 (ce qui correspond à $m = 3$).

EXERCICES

1. En utilisant l'exemple 2.5, démontrer que, si $0 \leq x \leq \pi$,

$$x(\pi - x) = \frac{8}{\pi} \sum_{m=1}^{\infty} \frac{\sin(2m-1)x}{(2m-1)^3}$$

Que donne cette relation au cas où $x = \pi/2$?
Appliquer la formule de Parseval au développement précédent et en déduire que

$$\frac{\pi^6}{960} = \sum_{m=1}^{\infty} \frac{1}{(2m-1)^6}$$

2. Soit f la fonction 2π-périodique égale à $(\pi - x)/2$ pour $0 < x < 2\pi$ et telle que $f(0) = f(2\pi) = 0$.
Calculer les coefficients de Fourier de f et les coefficients de Fourier de la primitive de f qui s'annule en 0.
En utilisant la formule de Parseval, montrer les formules suivantes

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \qquad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$$

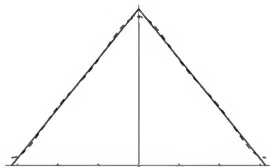


FIG. 2.3 –

Solution : Puisque f est impaire, sa série de Fourier ne contient que des termes en $\sin(nx)$. Une intégration par parties donne

$$b_n(f) = \frac{2}{\pi}\int_0^\pi \frac{1}{2}(\pi - x)\sin(nx)\,dx = \frac{1}{n}, \quad n \geq 1$$

La formule de Parseval dit que

$$\frac{1}{2}\sum_{n\geq 1}\frac{1}{n^2} = \frac{1}{2\pi}\int_0^{2\pi} |f(x)^2|\,dx = \frac{\pi^2}{12}$$

D'autre part, on vérifie facilement que la primitive de f qui s'annule en 0 et sa série de Fourier sont données par

$$\frac{\pi^2}{4} - \frac{(\pi - x)^2}{4} \sim \sum_{n\geq 1}\frac{1 - \cos nx}{n^2}$$

On en déduit que la série de Fourier de $x \mapsto (\pi - x)^2/4$ est

$$\frac{\pi^2}{12} + \sum_{n\geq 1}\frac{\cos nx}{n^2}$$

La deuxième relation résulte de la formule de Parseval. □

3. Soit $0 < h < \pi/2$ et soit f la fonction 2π-périodique , continue et paire définie par $f(0) = 1$, $f(x) = 0$ pour $2h \leq x \leq \pi$ et f est linéaire sur $[0, 2h]$. Déterminer la série de Fourier de f et écrire la formule de Parseval.
4. Montrer que chacune des suites $(\cos nx)$ et $(\sin nx)$, $n \in \mathbb{N}$, forme un système orthogonal et complet dans l'espace de Hilbert $L^2(0, \pi)$.
5. Soient f et g deux éléments de $L^2[0, 2\pi]$, montrer que $h = fg$ est dans $L^1[0, 2\pi]$ et que ses coefficients de Fourier sont donnés par par

$$c_n(h) = \sum_{k\in\mathbb{Z}} c_k(f)c_{n-k}(g), quadn \in \mathbb{Z}$$

où la série est absolument convergente.
Interpréter cette relation lorsque $n = 0$.
6. Soit $f \in L^2[0, 2\pi]$. Pour chaque entier relatif n, on pose

$$\gamma_n = \sum_{k\neq n} c_k(f)\frac{1}{n-k}$$

Montrer que la suite (γ_n) appartient à $\ell^2(\mathbb{Z})$ et que

$$\sum_{n\in\mathbb{Z}} |\gamma_n|^2 \leq \pi^2 \sum_{k\in\mathbb{Z}} |c_k(f)|^2$$

Les nombres γ_n sont l'analogue discret de $\int [f(t)/(x-t)]\, dt$.
Solution : Soit $g(x) = i(\pi - x)$. Compte tenu de l'exercice 2, on peut écrire

$$g(x) \sim \sum_{n\neq 0} \frac{e^{inx}}{n}, \quad \text{pour} \quad 0 < x < 2\pi$$

Les nombres γ_n sont les coefficients de Fourier de la fonction fg : celle-ci étant dans $L^2(0, 2\pi)$, la suite (γ_n) est dans $\ell^2(\mathbb{Z})$ et on a

$$\begin{aligned}\sum_{n\in\mathbb{Z}} |\gamma_n|^2 &= \frac{1}{2\pi}\int_0^{2\pi} |f(x)g(x)|^2\, dx \\ &\leq \frac{\pi^2}{2\pi}\int_0^{2\pi} |f(x)|^2\, dx = \pi^2 \sum_{k\in\mathbb{Z}} |c_k(f)|^2\end{aligned}$$

2.3 Polynômes de Chebyshev

L'application $\theta \mapsto \cos\theta$ est une bijection continue de $(0, \pi)$ sur $(-1, 1)$ et donc à chaque fonction F continue sur $(0, \pi)$ on associe de façon univoque la fonction f continue sur $(-1, 1)$ par la relation

$$F(\theta) = f(x) \quad \text{avec} \quad x = \cos\theta$$

De plus

$$\int_0^\pi |F(\theta)|^2 \, d\theta = \int_{-1}^1 |f(x)|^2 \, \frac{dx}{\sqrt{1-x^2}}$$

On en déduit, (voir le théorème de prolongement en annexe), que l'application qui à une fonction F de $C(0, \pi)$ fait corres-pondre la fonction f de $C(-1, 1)$, se prolonge en une bijection isométrique de l'espace de Hilbert $L^2((0, \pi), d\theta)$ sur l'espace de Hilbert $L^2((-1, 1), \omega(x)dx)$ des (classes de) fonctions de carré sommables sur $(-1, 1)$ pour la mesure de densité $\omega(x) = 1/\sqrt{1-x^2}$ par rapport à la mesure de Lebesgue, c'est-à-dire vérifiant

$$\|f\|_\omega^2 = \int_{-1}^1 |f(x)|^2 \frac{dx}{\sqrt{1-x^2}} < \infty$$

Le produit scalaire associé sera noté $\langle , \rangle_\omega$:

$$\langle f, g\rangle_\omega = \int_{-1}^1 \frac{f(x)\overline{g}(x)}{\sqrt{1-x^2}} dx, \quad \text{pour} \quad f, g \in L^2((-1, 1), \omega(x)dx)$$

On est en situation d'appliquer les résultats du paragraphe 1. La famille $\{1, x, \ldots, x^n, \ldots\}$ est totale dans l'espace $L^2((-1, 1), \omega(x)dx)$ et le procédé d'orthogonalisation de Gram-Schmidt permet d'en fabriquer une base hilbertienne. Cependant, l'isomorphisme mis en évidence plus haut permet de retrouver rapidement cette base hilbertienne, en effet les fonctions définies par $\phi_n(\theta) = \cos n\theta$ forment une base orthogonale de $L^2((0, \pi), d\theta)$ et on a

$$\int_0^\pi \phi_n(\theta)\overline{\phi_m(\theta)} \, d\theta = \begin{cases} (\pi/2)\delta_{nm}, & \text{si } n, m \neq 0\,; \\ \pi\delta_{0n}, & \text{sinon.} \end{cases}$$

Il en résulte que les fonctions $T_n(x) = \cos(n \arccos x)$ forment une base orthogonale de $L^2((-1, 1), \omega(x)dx)$ et on a

$$\langle T_n, T_m\rangle_\omega = \frac{\pi}{2}\delta_{nm} \quad \text{si} \quad n, m \neq 0$$
$$\langle T_n, T_0\rangle_\omega = \pi\delta_{0n} \qquad n \geq 0$$

On sait que $\cos n\theta$ s'exprime par un polynôme en $\cos\theta$. Il en résulte que $T_n(x)$ est en fait un polynôme de degré n en x, appelé polynôme de Chebyshev[6]. On a par exemple

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = \cos(2\theta) = 2x^2 - 1,$$
$$T_3(x) = \cos(3\theta) = 4x^3 - 3x, \quad T_4(x) = 8x^4 - 8x^2 + 1, \quad \text{etc...}$$

La figure 2.4 représente les graphes des polynômes T_n, pour $0 \leq n \leq 5$.
Le développement d'une fonction f de $L^2((-1,1), \omega(x)dx)$ suivant cette base s'écrit

$$\frac{1}{\pi}\langle f, T_0\rangle_\omega T_0 + \frac{2}{\pi}\sum_{n=1}^{\infty}\langle f, T_n\rangle_\omega T_n$$

où la série converge vers f dans $L^2((-1,1), \omega(x)dx)$. En vertu du théorème 1.4.7 (chapitre I), on peut énoncer :

Théorème 2.3.1. *La suite des polynômes de Chebyshev (T_n) est une base orthogonale de l'espace $L^2((-1,1), \omega(x)dx)$ et pour tout f dans cet espace, on a*

$$\lim_{n\to+\infty}\left\| f - \frac{1}{\pi}\langle f, T_0\rangle_\omega T_0 + \frac{2}{\pi}\sum_{k=1}^{n}\langle f, T_k\rangle_\omega T_k\right\|_\omega = 0$$

et

$$\|f\|_\omega^2 = \frac{1}{\pi}\left|\langle f, T_0\rangle\right|^2 + \frac{2}{\pi}\sum_{n=1}^{\infty}\left|\langle f, T_n\rangle\right|^2$$

On peut en déduire la propriété de minimisation suivante.

Corollaire 2.3.2. *Soit $f \in L^2((-1,1), \omega(x)dx)$. Dans l'espace vectoriel des polynômes P d'une variable réelle et de degré inférieur ou égal à $n \geq 0$, le polynôme*

$$T_n f = \frac{1}{\pi}\langle f, T_0\rangle_\omega T_0 + \frac{2}{\pi}\sum_{k=1}^{n}\langle f, T_k\rangle_\omega T_k$$

réalise le minimum de $\|f - P\|_\omega$, et il est le seul à avoir cette propriété.

Démonstration. Les polynômes T_n étant deux à deux orthogonaux, avec $\deg T_n = n$ pour tout n, on en déduit que $\{T_0, T_1, \dots T_n\}$ est une base de l'espace vectoriel des polynômes de degré inférieur ou égal à n. Le polynôme $T_n f$ n'est autre que la projection orthogonale de f sur cet espace vectoriel. □

[6]Le mathématicien russe Pafnoutiy Lvovitch CHEBYSHEV (1821-1894) a contribué à la Théorie Constructive des Fonctions et l'étude des Probabilités. Ses travaux sur la Théorie de l'Approximation restent des classiques de ce sujet.

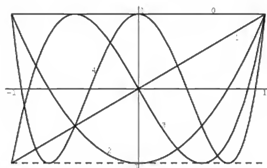


FIG. 2.4 –

Exemple 2.3.3. - Soit f la fonction définie pour x dans $(-1,1)$ par $f(x) = \sqrt{1-x^2}$. Considérons l'approximation de f par des sommes partielles de son développement suivant les polynômes de Chebyshev. Après changement de variables, on trouve $\langle f, T_0\rangle = 2$ et pour $n \geq 1$

$$\langle f, T_n\rangle_\omega = \int_0^\pi \sin\theta \cos n\theta \, d\theta = \begin{cases} 2/(1-n^2), & \text{si } n \text{ est pair ;} \\ 0, & \text{sinon.} \end{cases}$$

Le développement de f en série suivant les polynômes de Chebyshev est donc

$$\frac{2}{\pi} - \frac{4}{\pi}\sum_{n=1}^{\infty} \frac{1}{4n^2-1} T_{2n}(x)$$

Une application de la formule de Parseval donne dans ce cas

$$\frac{\pi^2}{8} = 1 + 2\sum_{n=1}^{\infty} \frac{1}{(4n^2-1)^2}$$

Les premières approximations de f dans $L^2((-1,1), \omega(x)dx)$ sont :

$$T_0 f(x) = \frac{2}{\pi}, \qquad T_2 f(x) = \frac{2}{3\pi}(5-4x^2),$$
$$T_4 f(x) = \frac{2}{15\pi}(23 - 4x^2 - 16x^4), \quad \text{etc...}$$

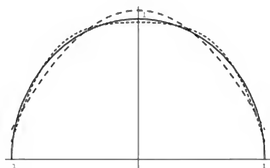


FIG. 2.5 –

Nous avons représenté sur la figure 2.5 les graphes de f, de $T_2 f$ et de $T_4 f$.

Exemple 2.3.4. - Considérons maintenant la fonction f définie par

$$f(x) = \begin{cases} -1, & \text{pour } -1 \leq x < 0\,; \\ 0, & \text{pour } x = 0\,; \\ 1, & \text{pour } 0 < x < 1. \end{cases}$$

Il est facile de voir que pour tout entier $n \geq 0$,

$$\langle f, T_{2n} \rangle = 0 \quad \text{et} \quad \langle f, T_{2n+1} \rangle = (-1)^n \frac{2}{2n+1}$$

Le développement en série de f suivant les polynômes de Chebyshev est donc

$$\frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} T_{2n+1}(x)$$

La figure 2.6 représente le graphe de f et les graphes de $T_5 f$ et $T_{11} f$ indiqués respectivement par 5 et 11. On remarque que dans l'exemple 3.3, le développement en série ne contient que les polynômes de Chebyshev d'indice pair, alors que celui de l'exemple 3.4 ne contient que ceux d'indice impair. Cela tient au fait que les fonctions considérées sont respectivement paire et impaire et au fait que T_n a la parité de n.

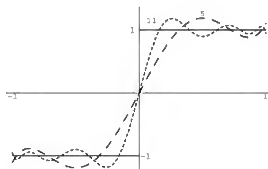


FIG. 2.6 –

A présent, nous allons dégager les principales propriétés des polynômes de Chebyshev, dont certaines ont un caractère général (voir le paragraphe 1 de ce chapitre).

Proriétés des Polynômes de Chebyshev

(1) *Parité.* Pour tout $n \in \mathbb{N}$, on a :

$$T_n(1) = 1 \quad \text{et} \quad T_n(-x) = (-1)^n T_n(x)$$

(2) *Relation de récurrence.* De l'identité trigonométrique

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta$$

on déduit que les polynômes de Chebyshev vérifient la relation de récurrence suivante

$$T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x)$$

Cette relation montre que le coefficient de x^n dans l'expression de $T_n(x)$ est 2^{n-1}. Elle permet aussi de calculer T_n par récurrence à partir de $T_0(x) = 1$ et $T_1(x) = x$.

(3) *Fonction génératrice.* On cherche une fonction de deux variables (x, z) telle que

$$G(x, z) = \sum_{n=0}^{\infty} T_n(x) z^n$$

En posant $x = \cos\theta$, on a

$$G(x,z) = \frac{1}{2}\sum_0^\infty \left(e^{im\theta} + e^{-im\theta}\right) z^m, \quad |z| < 1$$

$$= \frac{1}{2}\left[\frac{1}{1 - e^{i\theta} z} + \frac{1}{1 - e^{-i\theta} z}\right]$$

$$G(x,z) = \frac{1 - xz}{1 - 2xz + z^2}$$

(4) *Équation différentielle.* Pour tout $n \in \mathbb{N}$, le polynôme de Chebyshev T_n vérifie l'équation différentielle du second ordre suivante

$$(1 - x^2)T_n'' - xT_n' + n^2 T_n = 0$$

Pour le voir, il suffit de poser $x = \cos\theta$ et d'utiliser le fait que la fonction cosinus vérifie $\cos''(n\theta) + n^2\cos n\theta = 0$.

(5) *Racines et extremums de T_n.* Les racines de T_n se calculent facilement. En effet, puisque $\cos(nx) = 0$ si $x = (2k-1)\pi/(2n)$,

$$x_k = \cos\left(\frac{2k-1}{2n}\right)\pi, \quad \text{avec} \quad 1 \le k \le n$$

sont les n racines de T_n. Elles sont toutes réelles, simples, distinctes et appartiennent à l'intervalle $[-1, 1]$. D'autre part, la relation

$$T_n'(x) = \frac{n}{\sqrt{1-x^2}}\sin(n \arccos x)$$

montre que $T_n'(x) = 0$ si $x = x_k' = \cos(k\pi/n)$, où $1 \le k \le n-1$. On en déduit rapidement que T_n atteint ses extremums sur l'intervalle $[-1, 1]$, aux points x_k' et qu'en ces points

$$T_n(x_k') = \cos k\pi = (-1)^k, \quad 1 \le k \le n-1$$

de plus

$$T_n(x_0') = T_n(1) = 1 \quad \text{et} \quad T_n(x_n') = T_n(-1) = (-1)^n$$

On peut remarquer que les racines de T_n sont symétriques par rapport à 0 et que pour n grand, elles sont plus denses aux extrémités qu'au centre de l'intervalle $[-1, 1]$. Ces racines sont souvent utilisées comme points d'interpolation pour des fonctions continues sur $[-1, 1]$. □

(6) *Propriété de minimisation.* Les polynômes de Chebyshev jouent un rôle important en théorie de l'approximation. Cela tient à ce que, comme

l'a montré Chebyshev, ce sont là les polynômes s'écartant le moins de zéro sur le segment $[-1,1]$. Autrement dit, si l'on désigne par $\mathcal{P}_n$ l'ensemble des polynômes de degré n unitaires (c'est-à-dire dont le terme de plus haut degré est x^n), on a

$$\forall q \in \mathcal{P}_n, \qquad \sup_{|x|\leq 1} |q(x)| \geq \sup_{|x|\leq 1} \frac{1}{2^{n-1}} |T_n(x)| = \frac{1}{2^{n-1}}$$

EXERCICES

1. En se reportant aux expressions des polynômes T_0, T_2 et T_4, montrer par substitution que le monôme $f(x) = x^4$ s'écrit

$$x^4 = (3/8)T_0(x) + (1/2)T_2(x) + (1/8)T_4(x)$$

En déduire sans calcul les quantités $\langle f, T_n \rangle$, pour $0 \leq n \leq 4$. Trouver la combinaison linéaire $P = c_0T_0 + c_1T_1 + c_2T_2 + c_3T_3$ qui réalise la meilleure approximation en moyenne quadratique du monôme x^4 dans l'espace $L^2((-1,1),(1-x)^{-1/2}dx)$.

2. Soit f la fonction définie par$f(x) = |x|$. Montrer que $\langle f, T_0 \rangle = 2$ et que pour $n \geq 1$,

$$\langle f, T_n \rangle = \begin{cases} (-1)^k \frac{2}{4k^2-1}, & \text{si } n = 2k\,; \\ 0, & \text{si } n = 2k+1. \end{cases}$$

En déduire le développement en série de $|x|$ suivant les polynômes de Chebyshev et montrer que la convergence vers f a lieu non seulement dans l'espace de Hilbert $L^2((-1,1),\omega(x)dx)$, mais aussi uniformément sur $(-1,1)$.

3. Soit $f(x) = x^n$. Calculer le produit scalaire $\langle f, T_k \rangle$. Montrer que

$$x^n = \frac{1}{2^{n-1}} \left[T_n(x) + \binom{n}{1} T_{n-2} + \binom{n}{2} T_{n-4} + \cdots \right]$$

le dernier terme dépendant de la parité de n.
Solution. On part de la relation $(\cos\theta)^n = (e^{i\theta} + e^{-i\theta})^n$ et on développe le second membre grâce à la formule du binôme.

4. Démontrer la relation suivante (dite relation de Dirichlet)

$$\frac{1}{2} + T_1(x) + \cdots + \frac{1}{2} T_n(x) = \frac{T_n'(x)(1+x)}{2n}$$

Solution. Posons $x = \cos\theta$ et désignons le premier membre par $S(x)$,

il vient

$$
\begin{aligned}
S(x) &= \frac{1+\cos n\theta}{2} + \Re e\left(e^{i\theta}\frac{1-e^{i(n-1)\theta}}{1-e^{i\theta}}\right) \\
&= \frac{1+\cos n\theta}{2} + \frac{(\cos\theta - 1 - \cos n\theta + \cos(n-1)\theta)}{2(1-\cos\theta)} \\
&= \frac{\cos(n-1)\theta - \cos\theta\cos n\theta}{2(1-\cos\theta)} = \frac{(1+\cos\theta)\sin n\theta}{2\sin\theta}
\end{aligned}
$$

D'autre part, en tenant compte de la relation $x = \cos\theta$, on voit que $T_n'(x) = n\sin n\theta/\sin\theta$. La relation cherchée s'en déduit alors immédiatement. □

5. Soit $\overline{T}_n$ le polynôme donné par $\overline{T}_n = (1/2^{n-1})T_n$, $n \geq 1$.
(a) Montrer que $\overline{T}_n$ est un polynôme unitaire de degré n.
(b) Montrer que le maximum de $\mid \overline{T}_n(x) \mid$ sur $[-1,1]$ est égal à $1/2^{n-1}$ et que ce maximum est atteint $(n+1)$ fois aux points $x_k' = \cos k\pi/n$, $k = 0, 1\ldots, n$.
(c) Montrer que pour tout polynôme p unitaire et de degré n

$$\sup_{|x|\leq 1} |p(x)| \geq \sup_{|x|\leq 1} \frac{1}{2^{n-1}}$$

6. Les polynômes de Chebyshev de second espèce sont définis par

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}, \quad x = \cos\theta, \quad n = 0, 1\ldots$$

Démontrer que $U_n(x)$ est un polynôme de degré n et que l'on a

$$\int_{-1}^{1} U_m(x)U_n(x)\sqrt{1-x^2}\,dx = \frac{\pi}{2}\delta_{mn}$$

$$U_{n+1}(x) + U_{n-1}(x) = 2xU_n(x), n \geq 1.$$

Solution. Compte tenu du changement de variables $x = \cos\theta$, on peut voir que $T_n'(x) = n\sin n\theta/\sin\theta$. En développant le numérateur de l'expression définissant U_n, on voit que

$$U_n(x) = \frac{\sin n\theta}{\sin\theta}\cos\theta + \cos n\theta = \frac{1}{n}xT_n'(x) + T_n(x)$$

Il est alors évident que U_n est un polynôme de degré n. Les autres propriétés se démontrent facilement. □

2.4 Polynômes de Legendre

Prenons $I = (-1, 1)$ et $\omega = 1$. La famille $\{ x^n;\ 0 \leq n \}$ est totale dans l'espace de Hilbert $L^2(I; dx)$ et le procédé d'orthogonalisation de Gram-Schmidt donne

$$\begin{aligned}
&p_0(x) = 1, \\
&p_1(x) = x, \qquad \text{car} \quad \langle 1, x \rangle = 0 \\
&p_2(x) = x^2 - 1/3, \qquad \text{car} \quad \langle 1, x^2 \rangle = 1/3, \quad \|1\|^2 = 2 \quad \text{et} \quad \langle x, x^2 \rangle = 0 \\
&p_3(x) = x^3 - (3/5)x, \quad \text{car} \quad \langle x, x^3 \rangle = 2/5 \quad \text{et} \quad \|x\|^2 = 2/3
\end{aligned}$$

La formule suivante, dite *formule de Rodrigues*, donne l'expression de p_n pour tout entier $n \geq 0$.

Théorème 2.4.1. *Pour tout entier n, p_n est un polynôme unitaire de degré n et on a*

$$p_n(x) = \frac{n!}{(2n)!} \frac{d^n}{dx^n} \left((x^2 - 1)^n \right)$$

Démonstration. Il est clair que la formule de Rodrigues est vraie pour p_0 et p_1. D'autre part, comme $(x^2 - 1)^n$ est un polynôme de degré $2n$, sa dérivée d'ordre n est un polynôme de degré n, qui a la même parité que n et on vérifie facilement que le coefficient $n!/(2n)!$ a été choisi de façon que son terme de plus haut degré soit x^n. Compte tenu de l'unicité, il reste à démontrer que les polynômes définis par le second membre sont orthogonaux deux à deux, ce que l'on vérifie par intégrations par parties successives. En effet, pour deux entiers $n > m$,

$$\begin{aligned}
\langle p_n, p_m \rangle &= \frac{(n!)(m!)}{(2n)!(2m)!} \int_{-1}^{1} \frac{d^n}{dx^n} \left((x^2 - 1)^n \right) \frac{d^m}{dx^m} \left((x^2 - 1)^m \right) dx \\
&= (-1)^n \frac{(n!)(m!)}{(2n)!(2m)!} \int_{-1}^{1} (x^2 - 1) \frac{d^{n+m}}{dx^{n+m}} \left((x^2 - 1)^m \right) dx \\
&= 0, \quad \text{car} \quad n + m > 2m
\end{aligned}$$

La formule de Rodrigues est ainsi prouvée. □

Notons que pour tout n, $p_n(1) = 2^n (n!)^2 / (2n!)$.

Les polynômes de Legendre[7], que nous désignerons dans la suite par une

[7] Adrien Marie Legendre (1752-1833), mathématicien français né et mort à Paris. Le premier ouvrage qui rendit célèbre Adrien Marie Legendre a pour titre "Eléments de géométrie" (1794). Il représente un des premiers essais de formalisation rigoureuse de la géométrie, et il devait exercer une très grande influence sur les mathématiciens de son temps (vingt éditions de son vivant). Mais Legendre n'est pas uniquement connu comme géomètre et les domaines de ses recherches furent des plus variés : équations différentielles, calcul numérique, théorie des fonctions, théorie des nombres. Il a introduit les polynômes qui portent son nom en 1785, pour résoudre l'équation de Laplace en coordonnées sphériques.

lettre majuscule P_n, sont proportionnels aux polynômes p_n et normalisés de façon que $P_n(1)$ soit égal à 1. Ils sont donc donnés, grâce à la formule de Rodrigues, par

$$P_n(x) = \frac{1}{2^n(n!)}\frac{d^n}{dx^n}\Big((x^2-1)^n\Big) = \frac{(2n)!}{2^n(n!)^2}p_n(x), \quad n = 0, 1, \ldots$$

Ainsi, le coefficient de x^n dans l'expression de P_n est égal à $\frac{(2n)!}{2^n(n!)^2}$. On a par exemple

$$\begin{aligned} &P_0(x) = 1, \qquad P_1(x) = x, \quad P_2(x) = (1/2)(3x^2-1), \\ &P_3(x) = (1/2)(5x^3-3x), \qquad P_4(x) = (1/8)(35x^4-30x^2+3) \\ &P_5(x) = (1/8)(63x^5-70x^3+15x) \\ &P_6(x) = (1/16)(231x^6-315x^4+105x^2-5) \end{aligned}$$

La figure 2.7 représente les graphes de P_0, P_1, P_4, P_5 et P_6.

Théorème 2.4.2. *Pour tout entier $n \geq 0$, on a $\|P_n\|^2 = 2/(2n+1)$.*

Démonstration. En effet, un calcul simple d'intégrales donne

$$\int_{-1}^1 P_n(x)x^n\,dx = \frac{1}{2^n}\int_{-1}^1(1-x^2)^n\,dx = 2^{n+1}\frac{(n!)^2}{(2n+1)!}$$

Comme on sait que P_n est de la forme :

$$P_n(x) = \frac{(2n)!}{2^n(n!)^2}x^n + R_n(x)$$

où R_n est un polynôme de degré strictement plus petit que n, on en déduit que $\langle P_n, R_n\rangle = 0$ et par suite

$$\int_{-1}^1 P_n(x)^2\,dx = \frac{(2n)!}{2^n(n!)^2}\frac{2^{n+1}(n!)^2}{(2n+1)!} = \frac{2}{2n+1}$$

Le théorème 1.4.7 (chapitre I) s'énonce comme suit

Théorème 2.4.3. *La suite des polynômes de Legendre (P_n) est une base orthogonale de l'espace $L^2((-1,1),dx)$. Toute fonction f de $L^2((-1,1),dx)$ se développe de façon unique sous la forme*

$$f = \sum_{n=0}^{\infty} c_n(f)P_n \quad \textit{avec} \quad c_n(f) = \frac{2n+1}{2}\langle f, P_n\rangle$$

où la convergence de la série a lieu en moyenne quadratique

$$\lim_{N\to\infty}\left\|f - \sum_{n=0}^{N} c_n(f)P_n\right\| = 0$$

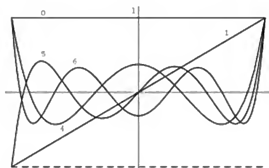


FIG. 2.7 –

On a de plus l'égalité de Parseval

$$\int_{-1}^{1} |f(x)|^2 \, dx = 2 \sum_{n=0}^{\infty} \frac{|c_n(f)|^2}{2n+1}$$

On notera que le polynôme de Legendre de f, de degré n, donné par

$$L_n f(x) = \sum_{k=0}^{n} c_k(f) P_k(x)$$

réalise la meilleure approximation en moyenne quadratique de f par des polynômes de degré inférieur ou égal à n.

On notera aussi que si f est paire (resp. impaire), son développement ne fera intervenir que les polynômes de Legendre d'indice pair (resp. impair).

Exemple 2.4.4. - Considérons la fonction $f(x) = |x|$. C'est une fonction paire. Son développement ne fera intervenir que les polynômes de Legendre d'indice pair et est donc de la forme $\sum_{n=0}^{\infty} c_{2n} P_{2n}$. En utilisant la parité et

la formule de Rodrigues, on a

$$\begin{aligned} c_{2n} &= \frac{(4n+1)}{2}\langle f, P_{2n}\rangle = (4n+1)\int_{-1}^{1}|x|P_{2n}(x)\,dx \\ &= \frac{(4n+1)}{2^{2n}(2n)!}\int_0^1 x\frac{d^{2n}}{dx^{2n}}((x^2-1)^{2n})\,dx \\ &= \frac{(4n+1)}{2^{2n}(2n)!}\int_0^1 \frac{d}{dx}\left[x\frac{d^{2n-1}}{dx^{2n-1}}(x^2-1)^{2n})\right]dx- \\ &\qquad - \frac{(4n+1)}{2^{2n}(2n)!}\int_0^1 \frac{d^{2n-1}}{dx^{2n-1}}(x^2-1)^{2n}dx \end{aligned}$$

La première intégrale du dernier membre est nulle alors que la deuxième donne

$$c_{2n} = \frac{(4n+1)}{2^{2n}(2n)!}\,\frac{d^{2n-2}}{dx^{2n-2}}(x^2-1)^{2n}\Big|_{x=0}$$

En développant $(x^2-1)^{2n}$ suivant la formule du binôme, on montre que le seul terme qui contribue dans le calcul de c_{2n} est celui pour lequel $k = n+1$, d'où

$$\begin{aligned} c_{2n} &= \frac{(4n+1)(-1)^{n+1}}{2^{2n}(2n)!}\,\frac{(2n)!(2n-2)!}{(n+1)!(n-1)!} \\ &= \frac{(4n+1)(-1)^{n+1}}{2^{2n}}\,\frac{(2n-2)!}{(n-1)!(n+1)!}, \quad n \geq 1, \quad c_0 = \frac{1}{2} \end{aligned}$$

Par exemple la meilleure approximation en moyenne quadratique de $|x|$ par des polynômes de degré inférieur ou égal à 6 est donnée par le polynôme

$$L_6 f(x) = \frac{1}{2} + \frac{5}{8}P_2(x) - \frac{3}{16}P_4(x) + \frac{13}{128}P_6(x)$$

La figure 2.8 représente le graphe de f et $L_6 f$.

Exemple 2.4.5. - Considérons la fonction f définie par

$$f(x) = \begin{cases} -1, & \text{pour } -1 \leq x < 0\,; \\ 0, & \text{pour } x = 0\,; \\ 1, & \text{pour } 0 < x < 1. \end{cases}$$

C'est une fonction impaire et son développement ne fera intervenir que les polynômes de Legendre d'indice impair. Calculons son polynôme de

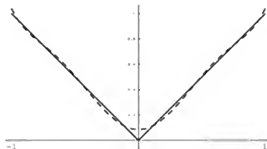


FIG. 2.8 –

Legendre de degré 5 :

$$c_1(f) = \frac{3}{2}\int_{-1}^{1} f(x)P_1(x)\,dx = 3\int_0^1 x\,dx = \frac{3}{2}$$
$$c_3(f) = \frac{7}{2}\int_{-1}^{1} f(x)P_3(x)\,dx = \frac{7}{2}\int_0^1 (5x^3 - 3x)\,dx = -\frac{3}{2}$$
$$c_5(f) = \frac{11}{2}\int_{-1}^{1} f(x)P_5(x)\,dx = \frac{11}{8}\int_0^1 (63x^5 - 70x^3 + 15x)\,dx = \frac{11}{16}$$

On en déduit que L_5 est donné par

$$L_5 f(x) = \frac{3}{2}P_1(x) - \frac{7}{8}P_3(x) + \frac{11}{16}P_5(x)$$

Sur la figure 2.9 on a représenté les graphes de f, de $L_5 f$ et de $L_7 f$.

Nous dégageons maintenant les principales propriétés des polynômes de Legendre.

Proriétés des polynômes de Legendre

(1) *Parité.* Pour tout entier naturel n et pour tout x, on a

$$P_n(-x) = (-1)^n P_n(x), \qquad \text{en particulier} \qquad P_n(-1) = (-1)^n$$

C'est une conséquence du corollaire 1.8. □

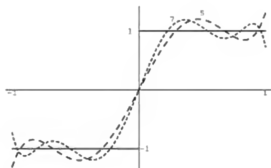


FIG. 2.9 –

Ainsi pour n pair, $P_n(x)$ ne contient que des puissances paires de x et pour n impair, $P_n(x)$ ne contient que des puissances impaires de x.

(2) *Représentation intégrale et majoration de P_n.* Pour tout entier n, on a

$$P_n(x) = \frac{1}{2\pi}\int_0^{2\pi}(x + i\sqrt{1-x^2}\sin\theta)^n\,d\theta,$$

en particulier $|P_n(x)| < 1$ pour tout x dans l'intervalle $]-1, 1[$.

Ce résultat, souvent employé dans les applications, est une conséquence de la formule de Rodrigues que l'on écrit, grâce à la formule de Cauchy, sous la forme

$$P_n(x) = \frac{1}{2^n}\frac{1}{2\pi i}\int_\gamma \frac{(z^2-1)^n}{(z-x)^{n+1}}\,dz$$

où γ désigne le cercle de centre x et de rayon $\sqrt{1-x^2}$. La représentation intégrale s'obtient alors en faisant, dans l'intégrale précédente, le changement de variable $z = x + i\sqrt{1-x^2}e^{i\theta}$. De plus, puisque

$$|x + i\sqrt{1-x^2}\sin\theta|^2 = \left[x^2 + (1-x^2)\sin^2\theta\right]^2 \leq 1$$

on en déduit la majoration $|P_n(x) \leq 1$, pour tout $-1 \leq x \leq 1$. □

(3) *Relation de récurrence.*

$$\frac{n+1}{2n+1}P_{n+1}(x) + \frac{n}{2n+1}P_{n-1}(x) = xP_n(x)$$

Pour démontrer cette relation on remarque que $xP_n(x)$ est un polynôme de degré $n+1$ et admet donc un développement sous la forme

$$xP_n(x) = \sum_{j\leq n+1} a_{nj}P_j(x) \qquad \text{avec} \qquad a_{nj}\|P_j\|^2 = \langle xP_n, P_j\rangle$$

Comme $a_{nj}\|P_j\|^2 = a_{jn}\|P_n\|^2$, il en résulte que $a_{nj} = 0$ pour $j < n-1$ (et pour $j > n+1$). En posant $a_{nn+1} = \alpha_n$, $a_{nn} = \beta_n$ et $a_{nn-1} = \gamma_n$, on en déduit que

$$xP_n(x) = \alpha_n P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x)$$

En examinant la parité des deux membres, on voit que $\beta_n = 0$ et en faisant $x = 1$, il vient $\alpha_n + \gamma_n = 1$. Enfin, en comparant les termes de plus haut degré dans chacun des deux membres, on trouve

$$\frac{(2n)!}{2^n(n!)^2} = \alpha_n \frac{(2n+2)!}{2^{n+1}((n+1)!)^2}$$

d'où il résulte que

$$\alpha_n = \frac{n+1}{2n+1} \quad \text{et par suite} \quad \gamma_n = \frac{n}{2n+1}$$

ce qui termine la preuve. □

(4) *Fonction génératrice.*

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n, \quad |t| < 1, |x| \leq 1$$

On désigne le premier membre par $u(x,t)$ et le second membre par $v(x,t)$. On a

$$u_t(x,t) = -\frac{1}{2}(1-2xt+t^2)^{-\frac{3}{2}}(-2x+2t) = (x-t)(1-2xt+t^2)^{-\frac{3}{2}}$$

où u_t désigne la dérivée partielle de u par rapport à t. Il en résulte que

$$(1-2xt+t^2)u_t(x,t) = (x-t)u(x,t)$$

D'autre part, puisque $|P_n(x)| \leq 1$, pour tout $n \in \mathbb{N}$ et tout $x \in [-1,1]$, on peut dériver par rapport à t terme à terme la série du second membre. Il vient

$$v_t(x,t) = \sum_{n=1}^{\infty} nP_n(x)t^{n-1}$$

On en déduit que

$$(1 - 2xt + t^2)v_t(x,t) = \sum_{n=1}^{\infty} nP_n(x)t^{n-1} - \sum_{n=1}^{\infty} 2xnP_n(x)t^n + \sum_{n=1}^{\infty} nP_n t^{n+1}$$

Le coefficient de t^n dans le second membre de l'égalité ci-dessus est $(n+1)P_{n+1}(x) - 2xnP_n(x) + (n-1)P_{n-1}(x)$, il est justement égal à $xP_n(x) - P_{n-1}(x)$ d'après la formule de récurrence. Par suite

$$(x-t)v(x,t) = (1 - 2xt + t^2)v_t(x,t)$$

Ainsi, les fonctions u et v vérifient, par rapport à t, la même équation différentielle (de degré 1) ; comme elles coïncident en $t = 0$, elles sont donc égales. □

(5) *Équation différentielle.* Pour tout entier naturel n, on a

$$(1 - x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0$$

On peut établir cette équation de la manière suivante ; on part de l'égalité

$$\frac{d}{dx}\{(1 - x^2)P_n'(x)\} = (1 - x^2)P_n''(x) - 2xP_n'(x)$$

Le deuxième membre est un polynôme de degré n, on peut donc le décomposer suivant la base $\{ P_i ;\ i \leq n \}$

$$\frac{d}{dx}\{(1 - x^2)P_n'(x)\} = \sum_{j=0}^{n} \alpha_i P_i(x) \qquad (\star)$$

La propriété d'orthogonalité implique

$$\alpha_i \|P_i\|^2 = \int_{-1}^{1} P_i(x) \frac{d}{dx}\{(1 - x^2)P_n'(x)\}dx$$

deux intégrations par parties successives donnent

$$\alpha_i \|P_i\|^2 = \int_{-1}^{1} P_n(x) \frac{d}{dx}\{(1 - x^2)P_i'(x)\}dx$$

Or la dérivée de $(1 - x^2)P_i'(x)$ est un polynôme de degré i, on en déduit que $\alpha_i = 0, \quad \forall\, i < n$ et la relation $(\star)$ peut maintenant s'écrire

$$(1 - x^2)P_n''(x) - 2xP_n'(x) = \alpha_n P_n(x)$$

En comparant les coefficients des termes de plus haut degré dans chacun des membres, on trouve immédiatement $\alpha_n = -n(n+1)$. L'équation différentielle cherchée s'en déduit.

(6) *Propriété de minimisation.* Les polynômes de Legendre sont caractérisés par la propriété de minimum suivante : Parmi tous les polynômes unitaires de degré n, p_n (défini au théorème 4.1) est l'unique qui réalise le minimum de la distance quadratique à 0, c'est-à-dire que pour tout polynôme unitaire q de degré n,

$$\int_{-1}^{1} |p_n(x)|^2\,dx \le \int_{-1}^{1} |q(x)|^2\,dx$$

et l'égalité a lieu si et seulement si $q = p_n$.

Cela résulte du fait que tout polynôme unitaire q de degré n s'écrit sous la forme $q(x) = p_n(x) + \sum_{k=0}^{n-1} a_k p_k$, donc

$$\int_{-1}^{1} |q(x)|^2\,dx = \int_{-1}^{1} |p_n(x)|^2\,dx + \sum_{k=0}^{n-1} |a_k|^2 \int_{-1}^{1} |p_k(x)|^2\,dx$$

Il est alors clair que le second membre atteint son minimum lorsque $a_k = 0$ pour $0 \le k \le n-1$. □

EXERCICES

1. Soit $f \in L^2(-1,1)$ et soit $f_N(x) = \sum_0^N a_n x^n$ telle que $\|f - f_N\|_2$ est minimale. Montrer que

$$\langle f, P_m \rangle = 0, \quad \forall m > N$$

2. Montrer que $\langle xP_n, P_n \rangle = 0$ pour tout entier n.

Solution. Comme P_n a la parité de n, le polynôme $xP_n^2(x)$ est impair et son intégrale sur $[-1,1]$ est donc nulle.

3. Trouver la constante a_n de façon que

$$x^n = a_n P_n(x) + Q_{n-2}(x)$$

où Q_{n-2} est un polynôme de degré inférieur ou égal à $n-2$.
En déduire que

$$\langle x^n, P_n \rangle = \frac{2^{n+1}(n!)^2}{(2n+1)!}, \quad \langle xP_{n-1}, P_n \rangle = \frac{2n}{(2n+1)(2n-1)}$$

Solution. L'expression de x^n résulte du fait que le coefficient du terme de plus haut degré dans l'expression de $P_n(x)$ est $(2n)!/(2^n(n!^2))$, et

des considérations de parité. Compte tenu de l'orthogonalité, on en déduit que

$$\langle x^n, P_n\rangle = \frac{2^n(n!)^2}{(2n)!}\|P_n\|^2 = \frac{2^{n+1}(n!)^2}{(2n+1)!}$$

La relation de récurrence que vérifient les polynômes de Legendre, permet d'écrire

$$\langle xP_{n-1}, P_n\rangle = \langle P_{n-1}, xP_n\rangle = \frac{n}{2n-1}\|P_n\|^2$$

Le résultat se déduit alors du théorème 4.2. □

4. Soit f une fonction de $L^2(-1,1)$ et désignons par $c_n(f)$ ses coefficients de Legendre. Montrer que

$$xf(x) = \sum_{n=0}^{\infty}\Big(\frac{n+1}{2n+3}c_{n+1}(f) + \frac{n}{2n-1}c_{n-1}(f)\Big)P_n(x)$$

5. Par intégration par parties, montrer que

$$\langle P'_{n+1}, x^m\rangle = 1 + (-1)^{m+n},\ \ m-1 < n+1$$

En déduire que

$$\langle P'_{n+1} - xP'_n, x^m\rangle = 0, \quad 0 \leq m < n$$

et conclure qu'il existe c_n, que l'on calculera, tel que

$$P'_{n+1} - xP'_n = c_nP_n$$

6. En utilisant l'expression de la fonction génératrice des polynômes de Legendre, montrer l'égalité suivante

$$\frac{2}{\sqrt{5-4x}} = \sum_0^{\infty}\frac{1}{2^n}P_n(x)$$

Solution. Il suffit de poser $t = 1/2$ dans l'expression de la fonction génératrice des polynômes de Legendre.

7. Montrer que les relations suivantes ont lieu pour tout $|t| < 1$:

$$\int_{-1}^{1}\frac{dx}{\sqrt{1-2xt+t^2}} = 2, \quad \int_{-1}^{1}\frac{xdx}{\sqrt{1-2xt+t^2}} = \frac{2}{3}t$$

$$\int_{-1}^{1}\frac{x^4dx}{\sqrt{1-2xt+t^2}} = \frac{1}{315}(16t^4 + 72t^2 + 126)$$

8. En utilisant la formule de récurrence vérifiée par les polynômes de Legendre, montrer la formule suivante dite de Christoffel-Darboux

$$\sum_{k=0}^{n}(2k+1)P_k(x)P_k(y) = (n+1)\frac{P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)}{x-y}$$

On pose $K_n(x,y) = \frac{1}{2}\sum_{k=0}^{n}(2k+1)P_k(x)P_k(y)$. Montrer que la fonction $y \mapsto K_n(x,y)$ est d'intégrale 1. Le noyau K_n est à rapprocher du noyau de Dirichlet. Montrer que pour toute fonction $f \in L^2((-1,1),dx)$, le polynôme de Legendre de degré n de f est donné par

$$L_n f(x) = \int_{-1}^{1} f(y)K_n(x,y)\,dy$$

2.5 Polynômes d'Hermite

Prenons $I = \mathbb{R}$, $\omega(x) = (2\pi)^{-\frac{1}{2}}\exp(-x^2/2)$ et soit $L^2(\mathbb{R},\omega(x)dx)$ l'espace des (classes de) fonctions de carré sommables pour la mesure de densité ω par rapport à la mesure de Lebesgue. Le produit scalaire et la norme seront notés respectivement $\langle,\rangle$ et $\|\ \|_2$. Pour f et g dans $L^2(\mathbb{R},\omega(x)dx)$,

$$\langle f,g\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)\overline{g(x)}\,e^{-x^2/2}dx$$

La densité ω a été choisie de façon que

$$\int_{-\infty}^{\infty}\omega(x)dx = 1$$

La famille $\{x^n; n \in \mathbb{N}\}$ est totale dans l'espace $L^2(\mathbb{R},\omega(x)dx)$ (théorème 2.1.5, chapitre II) et le procédé d'orthogonalisation de Gram-Schmidt fournit une base orthogonale formée de polynômes unitaires, appelés polynômes d'Hermite, que nous noterons dans la suite par (H_n).

Théorème 2.5.1. *Les polynômes d'Hermite*[8] *sont donnés par la formule de Rodrigues*

$$H_n(x) = (-1)^n e^{x^2/2}\frac{d^n}{dx^n}(e^{-x^2/2})$$

---

[8]Charles HERMITE (1822-1901), professeur d'Analyse à l'Ecole Polytechnique de 1869 à 1876, a été une des figures dominantes dans le développement de la Théorie des Formes Algébriques, de la Théorie Arithmétique des Formes Quadratiques et de la Théorie des fonctions elliptiques. Il a prouvé que le nombre e est transcendant et a donné la première solution de l'équation générale du cinquième degré grâce à l'utilisation de fonctions elliptiques.

Démonstration. Désignons par $Q_n(x)$ le second membre de la formule ci-dessus. Il s'écrit sous la forme

$$Q_n(x) = \frac{(-1)^n}{\omega(x)} \frac{d^n}{dx^n} \omega(x)$$

Il est facile de vérifier que pour tout n, Q_n est un polynôme unitaire de degré n. On montre maintenant l'orthogonalité $\langle Q_n, Q_m \rangle = 0$ lorsque $n \neq m$. Pour cela il suffit de montrer que $\langle x^n, Q_m \rangle = 0$ lorsque $n < m$. Or

$$\int_{-\infty}^{\infty} x^n Q_m(x)\omega(x)dx = \int_{-\infty}^{\infty} x^n (-1)^m \frac{d^m}{dx^m}(\omega(x))dx$$

Donc après n intégrations par parties, on obtient

$$\int_{-\infty}^{\infty} x^n Q_m(x)\omega(x)dx = (n!)(-1)^{m-n} \int_{-\infty}^{\infty} \frac{d^{m-n}}{dx^{m-n}}(\omega(x))\, dx$$

Si $n < m$ le second membre de cette égalité est nul, il en résulte que les polynômes (Q_n) sont deux à deux orthogonaux et compte tenu de l'unicité (voir théorème 1.6), on a bien $H_n = Q_n$. □

Si $n = m$, la dernière relation devient

$$\langle x^n, H_n \rangle = n! \int_{-\infty}^{\infty} \omega(x)dx = n!$$

ce qui permet de calculer la norme de H_n.

Corollaire 2.5.2. *Pour tout entier n,* $\|H_n\|^2 = n!$

Voici les sept premiers polynômes d'Hermite

$$H_0(x) = 1, \quad H_1(x) = x, \quad H_2(x) = x^2 - 1$$
$$H_3(x) = x^3 - 3x, \quad H_4(x) = x^4 - 6x^2 + 3,$$
$$H_5(x) = x^5 - 10x^3 + 15x, H_6(x) = x^6 - 15x^4 + 45x^2 - 15$$

En vertu du théorème 1.4.7 (chapitre 1), on peut énoncer

Théorème 2.5.3. *La suite (H_n) des polynômes d'Hermite est une base orthogonale de l'espace de Hilbert $L^2(\mathbb{R}, \omega(x)dx)$. Tout f dans $L^2(\mathbb{R}, \omega(x)dx)$ s'écrit*

$$f = \sum_{n=0}^{\infty} c_n(f) H_n \quad avec \quad c_n(f) = \frac{\langle f, H_n \rangle}{n!}$$

où la série converge vers f en moyenne quadratique

$$\lim_{N\to\infty} \int_{-\infty}^{\infty} \left| f(x) - \sum_{n=0}^{N} c_n(f) H_n(x) \right|^2 e^{-x^2/2}\, dx = 0$$

On a de plus l'égalité de Parseval $\|f\|_2^2 = \sum_{n=0}^{\infty} n!|c_n(f)|^2$.

Avant de traiter quelques exemples, dégageons d'abord les principales propriétés des polynômes d'Hermite.

Proriétés des polynômes d'Hermite

(1) *Parité.* Pour tout entier naturel n et pour tout x, on a

$$H_n(-x) = (-1)^n H_n(x)$$

C'est une conséquence du corollaire 1.8 et du fait que la densité ω est une fonction paire. □

Ainsi pour n pair, $H_n(x)$ ne contient que des puissances paires de x et pour n impair, $H_n(x)$ ne contient que des puissances impaires de x.

(2) *Relation de récurrence.* Les polynômes d'Hermite vérifient la relation de récurrence suivante, valable pour tout entier $n \geq 1$.

$$H_{n+1}(x) + nH_{n-1}(x) = xH_n(x)$$

Démonstration. Pour le voir, on part de l'identité suivante, dont la preuve ne présente aucune difficulté

$$\frac{d^n}{dx^n}(xF) = x\frac{d^n}{dx^n}F + n\frac{d^{n-1}}{dx^{n-1}}F$$

et on écrit

$$\begin{aligned} H_{n+1} &= \frac{(-1)^{n+1}}{\omega}\frac{d^{n+1}}{dx^{n+1}}(\omega) = \frac{(-1)^n}{\omega}\frac{d^n}{dx^n}(x\omega) \\ &= \frac{(-1)^n}{\omega}\left[x\frac{d^n}{dx^n}(\omega) + n\frac{d^{n-1}}{dx^{n-1}}(\omega)\right] \end{aligned}$$

ce qui est la relation cherchée. □

(3) *Fonction génératrice.* On a l'égalité suivante

$$G(x,t) = \exp(tx - t^2/2) = \sum_{n=0}^{\infty}\frac{t^n}{n!}H_n(x)$$

où la convergence a lieu dans $L^2(\mathbb{R}, \omega(x)dx)$.

Démonstration. On peut écrire $G(x,t) = \frac{1}{\omega(x)}\omega(t-x)$. La fonction ω est développable en série de Taylor convergente dans $\mathbb{R}$, il vient alors

$$G(x,t) = \frac{1}{\omega(x)}\sum_{n=0}^{\infty}\omega^{(n)}(-x)\frac{t^n}{n!}$$

Tenant compte de la définition de H_n et de la parité de ω, on en déduit la relation voulue. □

(4) *Équation différentielle.* Pour tout entier naturel n, on a

$$H_n''(x) - xH_n'(x) + nH_n(x) = 0$$

Démonstration. De la formule de Rodrigues on déduit l'égalité

$$H_n'(x) = xH_n(x) - H_{n+1}(x)$$

qui, comparée à la relation de récurrence, donne

$$H_n' = nH_{n-1}$$

En dérivant la relation de récurrence membre à membre et en tenant compte de cette dernière égalité (écrite avec $n+1$ à la place de n), on en déduit que H_n vérifie l'équation différentielle annoncée. □

Remarque 2.5.4. - L'équation différentielle vérifiée par H_n peut s'écrire sous la forme

$$L(H_n): = e^{x^2/2}\left(e^{-x^2/2}\frac{dH_n}{dx}\right)' = -nH_n$$

L'opérateur différentiel d'ordre deux, ainsi mis en évidence, est parfois appelé "opérateur d'Hermite". Son importance vient du fait que, si f et $L(f)$ sont dans $L^2(\mathbb{R}, e^{-x^2/2}dx)$, les coefficients du développement en série suivant la base (H_n), de $L(f)$ sont liés à ceux de f par l'égalité

$$c_n(Lf) = -nc_n(f)$$

Une application de cette remarque est donnée dans l'exercice 5.

Exemple 2.5.5. - Reprenons la fonction paire définie par $f(x) = |x|$. Son développement ne fera intervenir que des polynômes d'Hermite d'indice pair. On peut calculer facilement les premiers coefficients du développement de f, par exemple $c_0(f) = 2/\sqrt{2\pi}$, $c_2(f) = 1/\sqrt{2\pi}$ et

$$c_4(f) = -\frac{1}{12\sqrt{2\pi}}, \quad c_6(f) = \frac{1}{120\sqrt{2\pi}}$$

Un calcul direct et fastidieux, mais qui peut être évité (voir exercice 2), montre que pour tout $n \geq 1$,

$$H_{2n}(0) = \frac{(-1)^n(2n)!}{n!2^n}, \quad \text{d'où} \quad c_{2n}(f) = \frac{(-1)^{n-1}}{2^{n-1}(2n-1)(n!)\sqrt{2\pi}}$$

Nous avons représenté sur le figure 2.10 les graphes de f et de ses polynômes d'Hermite H_4f, H_6f et H_8f.

Exemple 2.5.6. - Considérons la fonction f définie par

$$f(x) = \begin{cases} -1, & \text{pour } x < 0\,; \\ 0, & \text{pour } x = 0\,; \\ 1, & \text{pour } 0 < x. \end{cases}$$

C'est une fonction impaire et son développement ne fera intervenir que les polynômes d'Hermite d'indice impair, c'est-à-dire que $c_{2n}(f) = 0$ pour tout entier n. Calculons $c_{2n+1}(f)$:

$$\begin{aligned} c_{2n+1}(f) &= \frac{\langle f, H_{2n+1}\rangle}{(2n+1)!} = \frac{-2}{(2n+1)!\sqrt{2\pi}} \int_0^\infty \frac{d^{2n+1}}{dx^{2n+1}}(e^{-x^2/2})\,dx \\ &= \frac{2}{(2n+1)!\sqrt{2\pi}} \frac{d^{2n}}{dx^{2n}}(e^{-x^2/2})\Big|_{x=0} = \frac{2}{(2n+1)!\sqrt{2\pi}} H_{2n}(0) \\ &= \frac{2}{(2n+1)!\sqrt{2\pi}} \frac{(-1)^n(2n)!}{n!2^n} = \frac{(-1)^n}{(2n+1)(n!)2^{n-1}\sqrt{2\pi}} \end{aligned}$$

Par exemple $c_1(f) = 2/\sqrt{2\pi}$, $c_3(f) = -1/(3\sqrt{2\pi})$, $c_5(f) = 1/(20\sqrt{2\pi})$. Les graphes de f et de ses polynômes d'Hermite $H_3 f$ et $H_5 f$ sont représentés sur la figure 2.11.

EXERCICES

1. Montrer les égalités suivantes :

$$\int_{-\infty}^{\infty} x^k \exp(-x^2/2)\,dx = \begin{cases} (2n)!/(n!2^n), & \text{si } k = 2n\,; \\ 0, & \text{sinon.} \end{cases}$$

2. En utilisant la fonction génératrice des polynômes d'Hermite, montrer que

$$\exp(-t^2/2) = \sum_{n=0}^{\infty} H_n(0) \frac{t^n}{n!}$$

En déduire que pour tout entier n, on a

$$H_{2n+1}(0) = 0, \qquad \text{et} \qquad H_{2n}(0) = \frac{(-1)^n(2n)!}{(n!)2^n}$$

Solution. En faisant $x = 0$ dans l'expression de la fonction génératrice des polynômes d'Hermite, on trouve la première relation . Comme son premier membre est une fonction paire, le second membre l'est aussi et donc $H_{2n+1}(0) = 0$, pour tout n. En posant $t^2 = s$, on en déduit que

$$e^{-s} = \sum_{n=0}^{\infty} H_{2n}(0) \frac{s^n}{(2n)!}$$

Le développement en série entière de la fonction e^{-s} permet d'obtenir imédiatement l'expression de $H_{2n}(0)$. □

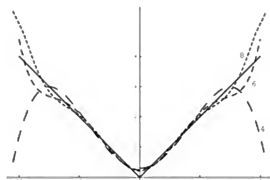


FIG. 2.10 –

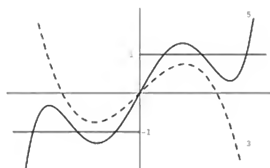


FIG. 2.11 –

3. A l'aide du procédé d'orthogonalisation de Gram-Schmidt, trouver les quatre premiers polynômes d'Hermite.
4. Montrer, par substitution, que

$$x^2 = H_0(x) + H_2(x)$$
$$x^3 = 3H_1(x) + H_3(x)$$
$$x^4 = 3H_0(x) + 6H_2(x) + H_4(x)$$

Déterminer le développement en série suivant les polynômes d'Hermite des fonctions $f(x) = x^{2r}$ et $g(x) = x^{2r+1}$, où $r \in \mathbb{N}$.
5. En utilisant la fonction génératrice des polynômes d'Hermite, montrer que

$$e^x = \sqrt{e}\sum_{n=0}^{\infty}\frac{H_n(x)}{n!}$$
$$(x-1)e^x = \sqrt{e}\sum_{n=1}^{\infty}\frac{H_n(x)}{(n-1)!}$$

Solution. En faisant $t = 0$ dans l'expression de la fonction génératrice des polynômes d'Hermite, on obtient la première relation. Pour la seconde relation, on considère l'opérateur différentiel défini par $L(u) = u'' - xu'$ qui sécrit sous la forme

$$L(u) = e^{x^2/2}\left(e^{-x^2/2}u'\right)'$$

L'équation différentielle que satisfait H_n s'écrit $L(H_n) = -nH_n$. Pour la fonction définie par $f(x) = e^x$, $Lf(x) = (1-x)f(x)$ appartient à l'espace $L^2(\mathbb{R}, e^{-x^2/2}dx)$ et par suite se développe dans la base (H_n). Deux intégrations par parties permettent de calculer les coefficients du développement de Lf :

$$\langle L(f), H_n\rangle = \langle f, L(H_n)\rangle = -n\langle f, H_n\rangle, \quad \text{pour tout } n$$

On en déduit immédiatement la relation cherchée. □
6. En utilisant la formule de Rodrigues et la relation de récurrence montrer les relations

$$H_n' = xH_n - H_{n+1} \quad \text{et} \quad H_n'(x) = nH_{n-1}(x)$$

En déduire (de deux manières) que

$$\langle H_n', H_m\rangle = \begin{cases} 0, & \text{si } m \neq n-1\,; \\ n!, & \text{si } m = n-1. \end{cases}$$

7. En utilisant l'exercice 6, montrer que si $f = \sum_0^\infty c_n H_n$, alors

$$xf = c_1 H_0 + \sum_{n=1}^{\infty} [c_{n-1} + (n+1)c_{n+1}]\, H_n$$

8. Écrire la formule de Parseval pour le développement de la fonction génératrice des polynômes d'Hermite.

9. Montrer la formule suivante dite de Christoffel-Darboux

$$\sum_{k=0}^{n} \frac{1}{k!} H_k(x) H_k(y) = \frac{1}{n!} \frac{H_{n+1}(x)H_n(y) - H_{n+1}(y)H_n(x)}{x - y}$$

On pose $K_n(x,y) = \sum_{k=0}^n 1/k! H_k(x) H_k(y)$, c'est l'analogue du noyau de Dirichlet. Montrer que

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} K_n(x,y)\, e^{-y^2/2} dy = 1$$

Soit f dans $L^2(\mathbb{R}, e^{-x^2/2}dx)$ et soit $H_n f$ le polynôme d'Hermite de degré n de f. Montrer que

$$H_n f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y) K_n(x,y) e^{-y^2/2}\, dy$$

2.6 Polynômes de Laguerre

Considérons l'espace $L^2((0,\infty), e^{-x}dx)$ des (classes de) fonctions de carré sommable pour la mesure de densité $\omega(x) = e^{-x}$ par rapport à la mesure de Lebesgue sur l'intervalle $(0, +\infty)$. Le produit scalaire et la norme sur $L^2((0,\infty), e^{-x}dx)$ sont donnés par

$$\langle f, g\rangle = \int_0^\infty f(x)\overline{g(x)} e^{-x}\, dx, \quad \|f\|_2 = \Big(\int_0^\infty |f(x)|^2 e^{-x}\, dx\Big)^{\frac{1}{2}}$$

L'espace de Hilbert $L^2((0,\infty), e^{-x}dx)$ est séparable, puisque la famille $\{1, x, \ldots, x^n, \ldots\}$ y est totale (théorème 2.1.5). Le procédé d'orthogonalisation de Gram-Schmidt appliqué à cette famille fournit une base orthogonale formée de polynômes $\tilde{L}_n$, avec $\deg \tilde{L}_n = n$. Les calculs dans ce cas sont particulièrement simples et utilisent l'égalité suivante

$$\int_0^\infty x^n e^{-x}\, dx = n!$$

Ainsi, on trouve par exemple que

$$\tilde{L}_0(x) = 1, \quad \tilde{L}_1(x) = x - 1, \quad \tilde{L}_2(x) = x^2 - 4x + 2,$$
$$\tilde{L}_3(x) = x^3 - 9x^2 + 18x - 6, \quad \tilde{L}_4(x) = x^4 - 16x^3 + 72x^2 - 96x + 24$$

Ce qui est convenu d'appeler polynômes de Laguerre[9] sont obtenus à partir des $\tilde{L}_n$ par la relation

$$L_n(x) = ((-1)^n/n!)\tilde{L}_n(x)$$

Les polynômes de Laguerre forment donc une base orthogonale de l'espace $L^2((0,\infty), e^{-x}dx)$. On peut les exprimer à l'aide de la formule de Rodrigues suivante

Théorème 2.6.1. *Pour tout entier n, on a :*

$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}(x^n e^{-x})$$

Démonstration. Le second membre de l'égalité ci-dessus est un polynôme de degré n, dont le terme de plus haut degré est $(-1)^n/(n!)$. Il suffit ensuite de montrer que ces polynômes sont deux à deux orthogonaux, ce qui résulte des égalités suivantes où $n > k$

$$\begin{aligned}\int_0^\infty x^k \frac{d^n}{dx^n}(x^n e^{-x})\,dx &= -k\int_0^\infty x^{k-1}\frac{d^{n-1}}{dx^{n-1}}(x^n e^{-x})\,dx \\ &= k(k-1)\int_0^\infty x^{k-2}\frac{d^{n-2}}{dx^{n-2}}(x^n e^{-x})\,dx \\ &= (-1)^k k!\int_0^\infty \frac{d^{n-k}}{dx^{n-k}}(x^n e^{-x})\,dx = 0\end{aligned}$$

Le théorème est ainsi prouvé. □

En appliquant la formule de Leibniz, on obtient le développement du polynôme L_n suivant les puissances de x

$$L_n(x) = \frac{1}{n!}\sum_{k=0}^{n}(-1)^k\binom{n}{k}n(n-1)\cdots(n-k+1)x^k$$

On peut remarquer alors que $L_n(0) = 1$ pour tout entier n.

[9]Edmon Nicolas Laguerre (1834–1886) est un mathématicien français.

Les calculs précédents permettent de trouver la norme de L_n, en effet

$$\begin{aligned}\|L_n\|^2 = \langle L_n, L_n\rangle &= \frac{(-1)^n}{(n!)}\langle x^n, L_n\rangle \\ &= \frac{(-1)^n}{(n!)^2}\int_0^\infty x^n \frac{d^n}{dx^n}(x^n e^{-x})\,dx \\ &= \frac{(-1)^{2n}}{n!}\int_0^\infty x^n e^{-x}\,dx = 1\end{aligned}$$

Par suite les polynômes de Laguerre (L_n) constituent une base hilbertienne de l'espace $L^2((0,\infty), e^{-x}dx)$.

Soit f un élément de $L^2((0,\infty), e^{-x}dx)$ et soit $c_n(f) = \langle f, L_n\rangle$ les coefficients de son développement suivant les polynômes de Laguerre. La formule de Parseval s'écrit

$$\int_0^\infty |f(x)|^2 e^{-x}\,dx = \sum_{n=0}^\infty |c_n(f)|^2$$

Exemple 2.6.2. - Soit $a > 0$ et soit f_a la fonction définie par

$$f_a(x) = e^{-ax}$$

Par des intégrations par parties, on a

$$\begin{aligned}c_n(f_a) &= \frac{1}{n!}\int_0^\infty e^{-ax}\frac{d^n}{dx^n}(x^n e^{-x})\,dx \\ &= \frac{a}{n!}\int_0^\infty e^{-ax}\frac{d^{n-1}}{dx^{n-1}}(x^n e^{-x})\,dx = \cdots \\ &= \frac{a^n}{n!}\int_0^\infty x^n e^{-(a+1)x}\,dx \\ &= \frac{a^n}{(a+1)^{n+1}}\end{aligned}$$

Nous en déduisons que

$$e^{-ax} = \frac{1}{1+a}\sum_{n=0}^\infty \Big(\frac{a}{1+a}\Big)^n L_n(x) \quad \text{pour} \quad x > 0$$

où la série converge au sens de la norme de $L^2((0,\infty), e^{-x}dx)$.
Les calculs qui suivent montrent qu'en fait, la convergence a lieu en tout point $x > 0$.

Fonction génératrice. On a

$$\frac{e^{-xt/(1-t)}}{1-t} = \sum_{n=0}^\infty L_n(x)t^n$$

où la série converge, non seulement dans $L^2((0,\infty), e^{-x}dx)$, mais aussi en tout point t, $|t| < 1$, et $x \geq 0$.

En effet, en utilisant le développement de $L_n(x)$ suivant les puissances de x (voir plus haut), on peut écrire

$$\begin{aligned}\sum_{n=0}^{\infty} L_n(x)t^n &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{x^k t^n}{k!}\\ &= \sum_{k=0}^{\infty}\frac{(-1)^k x^k}{k!}\sum_{n=k}^{\infty}\binom{n}{k}t^n\\ &= \sum_{k=0}^{\infty}\frac{(-1)^k x^k}{k!}\frac{t^k}{(1-t)^{k+1}} = \frac{e^{-xt/(1-t)}}{1-t}\end{aligned}$$

Les manipulations utilisées ici se justifient par la convergence absolue des séries qui interviennent. □

Remarque 2.6.3. - En faisant $t/(1-t) = a$ dans l'expression de la fonction génératrice, on retrouve la formule que nous avons établie à l'exemple 6.2 en utilisant la formule de Rodrigues. L'avantage que nous avons tiré est la convergence ponctuelle de la série. Cela étant, il faut bien noter que pour une fonction de $L^2((0,\infty), e^{-x}dx)$ quelconque, nous n'avons que la convergence en moyenne quadratique.

Relation de récurrence. Les polynômes de Laguerre vérifient la relation de récurrence suivante

$$(n+1)L_{n+1}(x) - (2n+1)L_n(x) + nL_{n-1}(x) = -xL_n(x)$$

Démonstration. Désignons par $G(x,t)$ la fonction génératrice des polynômes de Laguerre

$$G(x,t) = \frac{e^{-xt/(1-t)}}{1-t}$$

On vérifie facilement que $(1-t)^2 G'_t = (1-t-x)G$, ce qui donne

$$(1-2t+t^2)\sum_{n=0}^{\infty} nL_n(x)t^{n-1} = (1-t-x)\sum_{n=0}^{\infty} L_n(x)t^n$$

En identifiant les termes de même degré en t, on en déduit la relation de récurrence cherchée. □

Équation différentielle. Les polynômes de Laguerre vérifient l'équation différentielle du second ordre suivante :

$$xL_n''(x) + (1-x)L_n'(x) + nL_n = 0$$

Cette équation peut s'écrire sous la forme $\mathcal{L}(L_n) = nL_n$, où on a noté par $\mathcal{L}$ l'opérateur différentiel de Laguerre défini, pour une fonction u dans $C^2(\mathbb{R})$, par $\mathcal{L}(u) = e^x(xe^{-x}u')'$.

Démonstration. En effet, $P(x) = xL_n''(x) + (1-x)L_n'(x)$ étant un polynôme de degré n, il s'écrit sous la forme

$$P(x) = \sum_{k=0}^{n} c_k L_k(x), \quad \text{où} \quad c_k = \langle P, L_k \rangle$$

Pour calculer les coefficients c_k, on note que $P(x) = e^x(xe^{-x}L_n')'$, et une intégration par parties donne

$$c_k = -\int_0^\infty L_k'(x)L_n'(x)xe^{-x}\,dx$$

Or les polynômes (L_n') sont deux à deux orthogonaux dans $L^2((0,\infty), xe^{-x})$ (voir exercice 2), on en déduit que $c_k = 0$ pour tout $k \neq n$ et par suite $P(x) = c_n L_n(x)$. En identifiant les termes de plus haut degré des deux membres, on trouve $c_n = -n$. □

EXERCICES

1. Montrer, à l'aide de la formule de Rodrigues, que la relation suivante est satisfaite pour tout entier r

$$x^r = (r!)^2 \sum_{n=0}^{r} \frac{(-1)^n}{n!(r-n)!} L_n(x)$$

2. Montrer que pour tout $m < n$

$$\int_0^\infty x^{m-1}(1-x)L_n(x)e^{-x}\,dx = 0$$

En remarquant que $(1-x)e^{-x} = (xe^{-x})'$, montrer que

$$\int_0^\infty x^{m-1}L_n'(x)xe^{-x}\,dx + (m-1)\int_0^\infty x^{m-2}L_n(x)e^{-x}\,dx = 0$$

Montrer que la deuxième intégrale est nulle et en déduire que

$$\int_0^\infty x^{m-1}L_n'(x)xe^{-x}\,dx = 0, \quad \text{pour} \quad m < n$$

En déduire que les polynômes (L_n') forment une base orthogonale de l'espace de Hilbert $L^2((0,\infty), xe^{-x}dx)$.

Solution. Comme $(1-x)x^{m-1}$ est un polynôme de degré m, il est orthogonal à $L_n(x)$ pour tout $n > m$, ce qu'on peut traduire par la relation

$$\int_0^\infty (xe^{-x})'x^{m-1}L_n(x)\,dx = 0, \quad \text{pour} \quad n > m$$

Faisons une intégration par parties, il vient

$$\int_0^\infty x^{m-1} L_n'(x) x e^{-x}\, dx + (m-1) \int_0^\infty x^{m-2} L_n(x) e^{-x}\, dx = 0$$

La deuxième intégrale est nulle parce que les polynômes x^{m-2} et $L_n(x)$ sont orthogonaux et il s'en suit que pour $n > m$,

$$\int_0^\infty x^{m-1} L_n'(x) x e^{-x}\, dx = 0$$

Comme pour tout n, $L_n'(x)$ est un polynôme de degré $n-1$, la relation précédente traduit le fait que la suite (L_n') est une base orthogonale de $L^2((0,\infty), xe^{-x}dx)$.

3. Montrer, en utilisant la formule de Rodrigues et la formule de Cauchy, que

$$L_n(x) = \frac{e^x}{2i\pi} \int_{\mathcal{C}_x} \frac{t^n e^{-t}}{(t-x)^{n+1}}\, dt$$

où $\mathcal{C}_x$ est un cercle centré en $x > 0$ et de rayon suffisamment petit. En faisant le changement de variable $z = \frac{t-x}{t}$, montrer que l'on a aussi

$$L_n(x) = \frac{1}{2i\pi} \int_{\mathcal{C}_0} \frac{e^{-xz/(1-z)}}{(1-z) z^{n+1}}\, dz$$

Interpréter le second membre de cette égalité à l'aide du théorème de Cauchy et retrouver la relation

$$\frac{e^{-xz/(1-z)}}{(1-z)} = \sum_{n=0}^\infty L_n(x) z^n, \quad |z| < 1$$

4. On désigne par G la fonction génératrice des polynômes de Laguerre. Montrer les égalités suivantes, où N est un entier positif.

$$\int_0^\infty \left| G(x,t) - \sum_{n=0}^N L_n(x) t^n \right|^2 e^{-x}\, dx = \frac{1}{1-t^2} - \sum_{n=0}^N t^{2n}$$

$$\int_0^\infty \left| G(x,t) \right|^2 e^{-x}\, dx = \frac{1}{1-t^2}$$

Solution. C'est une application directe de la formule de Parseval et de l'égalité évidente $1/(1-t^2) = \sum_0^\infty t^{2n}$, valable pour $|t| < 1$. Bien sûr, la deuxième égalité peut se démontrer par un calcul direct très simple.

5. Pour $\alpha > -1$, on définit les polynômes de Laguerre L_n^α, d'indice α, par

$$L_n^\alpha(x) = \frac{x^\alpha e^x}{n!} \frac{d^n}{dx^n}(x^{n-\alpha} e^{-x})$$

(a) Montrer que L_n^α est un polynôme de degré n, dont le coefficient du terme de plus haut degré est $(-1)^n/(n!)$.
(b) Montrer que

$$\int_0^\infty L_n^\alpha(x) L_m^\alpha(x) x^\alpha e^{-x}\, dx = \frac{\Gamma(n+\alpha+1)}{n!}\delta_{nm}$$

et en déduire que les polynômes (L_n^α) forment une base orthogonale de l'espace de Hilbert $L^2((0,\infty); x^\alpha e^{-x} dx)$.
(c) Montrer que le polynôme L_n^α vérifie une équation différentielle que l'on explicitera.
(d) Soit $a > 0$ et $f_a(x) = e^{-ax}$. Montrer que f_a appartient à l'espace $L^2((0,\infty); x^\alpha e^{-x} dx)$ et que son développement suivant la base (L_n^α) est donné par

$$e^{-ax} = \sum_{n=0}^\infty \frac{a^n}{(a+1)^{n+\alpha+1}} L_n^\alpha(x)$$

(e) En déduire la relation

$$(1-t)^{-\alpha-1} e^{-xt/(1-t)} = \sum_{n=0}^\infty L_n^\alpha(x) t^n,$$

valable quel que soit t, $|t| < 1$.

Chapitre 3

Endomorphismes continus d'un espace de Hilbert

L'étude des opérateurs linéaires continus occupe une partie importante en analyse hilbertienne. Dans ce chapitre on donne les propriétés générales des opérateurs linéaires continus dans un espace de Hilbert : inversibilité, spectre, adjoint, opérateurs auto-adjoints, opérateurs unitaires, opérateurs de projection orthogonale.
Rappelons d'abord quelques propriétés élémentaires valables dans le cadre des espaces vectoriels normés sur $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

3.1 Généralités sur les opérateurs continus

Proposition 3.1.1. *Soient E et F deux espaces vectoriels normés sur $\mathbb{K}$ et A un opérateur linéaire de E dans F. Il y a équivalence entre :*

(a) *l'opérateur linéaire A est continu*

(b) *l'opérateur linéaire A est continu en 0.*

(c) *l'opérateur linéaire A est continu en un point.*

(d) *Il existe une constante $c > 0$ telle que $\|Ax\| \leq c\|x\|$, pour $x \in E$.*

Démonstration. Il est clair que $(a) \Longrightarrow (b) \Longrightarrow (c)$ et $(d) \Longrightarrow (b)$. D'autre part la démonstration de $(b) \Longrightarrow (d)$ est en tout point identique à celle de la proposition 3.2 (chapitre I). Il reste à prouver que $(c) \Longrightarrow (a)$: Supposons A continu en x_0 et soit $x \in E$ et (x_n) une suite qui converge vers x. La suite $(x_n - x + x_0)$ converge vers x_0 donc $A(x_n - x + x_0)$ converge vers Ax_0. Comme A est linéaire on en déduit immédiatement que Ax_n converge vers Ax. □

Notons que l'assertion (d) traduit le fait que A est un opérateur linéaire borné. Ainsi un opérateur linéaire de E dans F est continu si, et seulement si, il est borné.

Notations. On notee par $\mathcal{L}(E, F)$ l'ensemble des opérateurs linéaires continus de E dans F. Lorsque $F = \mathbb{K}$, $\mathcal{L}(E, \mathbb{K})$ est le dual topologique de E, il est noté E'. Enfin, si $E = F$ on écrira $\mathcal{L}(E)$ au lieu de $\mathcal{L}(E, E)$. Pour un élément A de $\mathcal{L}(E, F)$, on pose

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\|$$

Proposition 3.1.2. *Soient E et F deux espaces vectoriels normés.*

(a) *Si A et B sont dans $\mathcal{L}(E, F)$, alors $A + B$ est dans $\mathcal{L}(E, F)$ et*

$$\|A + B\| \leq \|A\| + \|B\|$$

(b) *Si $\alpha \in \mathbb{K}$ et $A \in \mathcal{L}(E, F)$, alors αA est dans $\mathcal{L}(E, F)$ et*

$$\|\alpha A\| = |\alpha| \, \|A\|$$

(c) *Si A et B sont dans $\mathcal{L}(E)$, leur composé AB est dans $\mathcal{L}(E)$ et*

$$\|AB\| \leq \|A\| \, \|B\| \quad \text{et} \quad \|A^n\| \leq \|A\|^n, \quad \textit{avec} \quad n \in \mathbb{N}$$

De cette proposition, dont la preuve est évidente, on déduit que l'application qui à un élément A de $\mathcal{L}(E, F)$ associe $\|A\|$ est une norme qui fait de $\mathcal{L}(E, F)$ un espace vectoriel normé.

Notons que cette norme peut se définir encore par les relations

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\| = 1} \|Ax\|$$
$$= \inf\Big\{c > 0 \mid \|Ax\| \leq c\|x\|; \ x \in E\Big\}$$

Démonstration. En utilisant la linéarité, on démontre facilement les deux premières relations. Pour la dernière, posons

$$\alpha = \inf\{ c > 0 \,/\, \|Ax\| \leq c\|x\|, \ x \in E\}$$

Pour tout ϵ strictement positif, il vient $\big\|A\big((\|x\| + \epsilon)^{-1}x\big)\big\| \leq \|A\|$, donc pour tout x dans E, on a $\|Ax\| \leq \|A\|(\|x\| + \epsilon)$ et par suite $\|Ax\| \leq \|A\| \, \|x\|$, ce qui se traduit par l'inégalité $\alpha \leq \|A\|$. D'un autre côté, si $\|Ax\| \leq c\|x\|$ pour tout $x \in E$, alors $\|A\| \leq c$, et par suite $\|A\| \leq \alpha$, d'où résulte l'égalité $\|A\| = \alpha$. Ainsi $\|Ax\| \leq \|A\| \, \|x\|$, pour tout $x \in E$ et $\|A\|$ est la plus petite constante qui réalise cette inégalité. □

Notons que la convergence dans $\mathcal{L}(E, F)$ d'une suite (A_n) vers A signifie que $\|A - A_n\|$ tend vers 0, lorsque n tend vers l'infini.

Proposition 3.1.3. *Si F est complet alors $\mathcal{L}(E, F)$ est complet.*

Démonstration. Supposons F un espace de Banach et soit (A_n) une suite de Cauchy de $\mathcal{L}(E, F)$:

$$\forall \epsilon > 0, \exists N \mid n \geq N \text{ et } m \geq N \Longrightarrow \|A_n - A_m\| \leq \epsilon$$

On en déduit que, pour tout élément x de E, et $n, m \geq N$, on a

$$\|A_n x - A_m x\| \leq \|A_n - A_m\| \, \|x\| \leq \epsilon \|x\|$$

de sorte que les éléments $A_n x$ forment une suite de Cauchy dans F ; celui-ci étant complet, cette suite converge vers un élément y de F. Posons $y = Ax$; on vérifie facilement que A est linéaire, de plus, par passage à la limite quand m tend vers l'infini dans l'inégalité précédente, on obtient

$$n \geq N \Longrightarrow \|A_n x - Ax\| \leq \epsilon \|x\|$$

Comme A_n est continu, cette inégalité montre que A est continu et que, pour $n \geq N$, on a $\|A_n - A\| \leq \epsilon$. Il en résulte que la suite (A_n) converge en norme vers A. □

Remarque 3.1.4. -

• Il arrive qu'une suite (A_n) d'éléments de $\mathcal{L}(E, F)$ possède la propriété suivante : $(A_n x)$ converge vers Ax, pour tout x dans E, mais $\|A - A_n\|$ ne tend pas vers 0, quand n tend vers l'infini. En voici un exemple simple : sur l'espace $\ell^2(\mathbb{N})$, muni de sa base canonique, on considère les opérateurs A_n définis par

$$A_n e_k = \begin{cases} e_1, & \text{si } k = n \\ 0, & \text{si } k \neq n \end{cases}$$

Cela veut dire que pour tout $\mathbf{x}$ dans $\ell^2(\mathbb{N})$, $A_n \mathbf{x} = x_n e_1$. Il est clair que $\|A_n\| = 1$, pour tout n, alors que la limite de $A_n \mathbf{x}$, quand n tend vers l'infini, est nulle pour tout élément $\mathbf{x}$ de $\ell^2(\mathbb{N})$.

• Insistons sur le fait suivant, qui n'est autre qu'une traduction du théorème de Banach-Steinhaus (voir Annexe) :

(a) *Si une suite (A_n) de $\mathcal{L}(E, F)$ est telle que, pour tout $x \in E$, la suite $(A_n x)$ est bornée, alors il existe une constante M telle que, pour tout n, $\|A_n\| \leq M$.*

(b) *Si une suite (A_n) de $\mathcal{L}(E, F)$ est telle que, pour tout $x \in E$, la suite $(A_n x)$ a une limite dans F, alors l'application A, qui à x associe $\lim_{n \to \infty} A_n x$, est un opérateur linéaire continu de E dans F.*

Nous allons voir maintenant que, si l'espace de Hilbert E est séparable, tout élément de $\mathcal{L}(E)$ admet, dans une base hilbertienne donnée, une "représentation matricielle" tout à fait analogue à celle, bien connue, lorsque E est de dimension finie.

Soient (e_n) une base hilbertienne de E et A un élément de $\mathcal{L}(E)$. Tout $x \in E$ s'écrit de manière unique

$$x = \sum_{1}^{\infty} x_j e_j \quad \text{avec} \quad \sum_{j=1}^{\infty} |x_j|^2 < \infty$$

Comme A est continu, il vient $Ax = \sum_{j=1}^{\infty} x_j A e_j$. L'opérateur A est ainsi bien déterminé par son action sur les vecteurs de base et les composantes (y_i) de Ax, dans la base (e_i), sont données par

$$y_i = \langle Ax, e_i \rangle = \sum_{j=1}^{\infty} x_j \langle A e_j, e_i \rangle$$

Si l'on pose $a_{ij} = \langle A e_j, e_i \rangle$, on aura

$$A e_j = \sum_{i=1}^{\infty} a_{ij} e_i, \qquad \text{avec} \quad \|A e_j\|^2 = \sum_{i=1}^{\infty} |a_{ij}|^2$$

et

$$Ax = \sum_{i=1}^{\infty} y_i e_i, \quad \text{avec} \quad y_i = \sum_{j=1}^{\infty} a_{ij} x_j$$

Ainsi, dans $\ell^2(\mathbb{N})$ muni de la base canonique, A est représenté par l'application qui à la suite (x_j) fait correspondre la suite (y_i), avec $y_i = \sum_{j=1}^{\infty} a_{ij} x_j$. L'opérateur A est donc représenté par la *matrice infinie* (a_{ij}), $i, j \in \mathbb{N}$.

EXERCICES

1. Soit $L^2(0,1)$ l'espace de Hilbert des fonctions de carré intégrables sur $(0,1)$, relativement à la mesure de Lebesgue. Soit l'opérateur défini sur $L^2(0,1)$ par $Af = xf$.
Montrer que A est un opérateur linéaire continu et calculer sa norme.

Solution : Il est clair que l'opérateur défini sur $L^2(0,1)$ par $Af = xf$ est linéaire. D'autre part, pour tout f dans $L^2(0,1)$, on a :

$$\|Af\|^2 = \int_0^1 x^2 |f(x)|^2 dx \leq \int_0^1 |f(x)|^2 dx = \|f\|^2$$

Cela montre que l'opérateur A est continu sur $L^2(0,1)$ et que $\|A\| \leq 1$. Pour $n \in \mathbb{N}^*$, on désigne par f_n la fonction caractéristique de l'intervalle $[1 - 1/n, 1]$. On a

$$\|Af_n\|^2 = \int_{1-1/n}^{1} x^2 dx \geq \int_{1-1/n}^{1} \left(1 - \frac{1}{n}\right)^2 dx = \left(1 - \frac{1}{n}\right)^2 \|f_n\|^2$$

Il en résulte que, pour tout entier $n > 0$, on a

$$\frac{\|Af_n\|}{\|f_n\|} \geq 1 - \frac{1}{n}$$

On en déduit que $\|A\| = 1$. □

2. Soit E un espace de Hilbert sur $\mathbb{C}$ et soit $a(,)$ une forme sesquilinéaire et continue sur $E \times E$, c'est-à-dire qu'il existe une constante M telle que : $|a(x,y)| \leq M\|x\|\|y\|$, pour $x, y \in E$.
En utilisant le théorème de représentation de Riesz, montrer qu'il existe $A \in \mathcal{L}(E)$ tel que, pour tout x et tout y dans E, on ait

$$\langle x, Ay \rangle = a(x,y)$$

Solution : Les hypothèses montrent que, pour tout y, l'application $x \mapsto a(x,y)$ est une forme linéaire continue. D'après le théorème de représentation de Riesz, il existe un unique élément qui dépend de y et que nous notons $w(y)$ tel que

$$a(x,y) = \langle x, w(y) \rangle, \quad \forall x \in E$$

et tel que $\|w(y)\| \leq M\|y\|$. Cette inégalité montre que l'application qui à y associe $w(y)$ est continue. On vérifie qu'elle est linéaire :

$$\forall x \in E, \quad a(x, y+z) = \langle x, w(y+z) \rangle$$

Comme a est sesquilinéaire, on a aussi

$$\forall x \in E, \quad a(x, y+z) = a(x,y) + a(x,z) = \langle x, w(y) \rangle + \langle x, w(z) \rangle$$

c'est-à-dire

$$\forall x \in E, \quad \langle x, w(y+z) \rangle = \langle x, w(y) \rangle + \langle x, w(z) \rangle$$

On en déduit que $w(y+z) = w(y) + w(z)$. D'autre part, pour tout scalaire α, on a

$$a(x, \alpha y) = \overline{\alpha} a(x,y)$$

Le premier membre de cette égalité vaut $\langle x, w(\alpha y) \rangle$ et le second vaut $\overline{\alpha}\langle x, w(y) \rangle = \langle x, \alpha w(y) \rangle$. Il en résulte que $w(\alpha y) = \alpha w(y)$. Ainsi, on a prouvé que l'application $y \mapsto w(y)$ est linéaire et continue. Il suffit donc de poser $w(y) = Ay$. □

3. Montrer que si (A_n) est une suite de Cauchy d'éléments de $\mathcal{L}(E)$, alors $(\|A_n\|)$ est elle aussi une suite de Cauchy.

Solution : L'inégalité triangulaire montre que, pour deux éléments A et B de $\mathcal{L}(E)$, on a $\|A\| \leq \|A - B\| + \|B\|$. On en déduit alors l'inégalité $\Big|\|A_n\| - \|A_m\|\Big| \leq \|A_n - A_m\|$. Par conséquent, si (A_n) est une suite de Cauchy, il en sera de même de la suite $(\|A_n\|)$. □

4. Soit (A_n) une suite d'éléments de $\mathcal{L}(E)$ qui converge vers A et soit (x_n) une suite d'éléments de E qui converge vers x. Montrer que la suite $(A_n x_n)$ converge vers Ax.

Solution : L'inégalité triangulaire permet d'écrire

$$\|Ax - A_n x_n\| \leq \|A\| \, \|x - x_n\| + \|x_n\| \, \|A - A_n\|$$

Lorsque n tend vers l'infini, le premier terme du second membre tend vers 0 et le second terme fait de même car la suite (x_n), étant convergente, est bornée et par hypothèse $\|A - A_n\| \to 0$ $(n \to \infty)$. □

5. Soient E un espace de Hilbert et (A_n) une suite d'éléments de $\mathcal{L}(E)$. On suppose qu'il existe une constante $C > 0$ telle que $\|A_n\| \leq C$ pour tout entier n. On suppose qu'il existe une partie F dense dans E telle que, pour tout $x \in F$, la suite $(A_n x)$ possède une limite. Montrer que, pour tout $x \in E$, la suite $(A_n x)$ a une limite et que l'application $x \mapsto \lim_{n \to \infty} A_n x$ définit un élément de $\mathcal{L}(E)$.

Solution : Soit x un élément de E, on va montrer que la suite $(A_n x)$ est une suite de Cauchy. Puisque F est dense dans E, il existe une suite (x_p) d'éléments de F qui converge vers x. Pour tous les entiers n, m et p, on a

$$\|A_n x - A_m x\| \leq \|A_n x - A_n x_p\| + \|A_n x_p - A_m x_p\| + \|A_m x_p - A_m x\|$$

$$\leq 2C\|x - x_p\| + \|A_n x_p - A_m x_p\|$$

Pour tout $\epsilon > 0$, il existe un entier p tel que $\|x - x_p\| \leq \epsilon/2C$. Puisque x_p appartient à F, la suite $n \mapsto A_n x_p$ est convergente. Il existe donc un entier N tel que

$$\forall n, m \geq N, \quad \|A_n x_p - A_m x_p\| \leq \epsilon$$

On en déduit que, pour $n, m \geq N$, $\|A_n x - A_m x\| \leq 2\epsilon$; (Ax_n) est donc une suite de Cauchy. Si l'on désigne par Ax sa limite, on vérifie que $x \mapsto Ax$ est linéaire, de plus, pour tout entier $n \geq N$, on a

$$\|Ax\| \leq \|Ax - A_n x\| + \|A_n x\| \leq \epsilon + C\|x\|$$

Il en résulte que l'opérateur A est continu et que $\|A\| \leq C$. □

6. Sur l'espace de Hilbert $\ell^2(\mathbb{N})$, muni de sa base canonique (e_n), on définit les opérateurs A_n par

$$A_n e_k = \begin{cases} e_n, & \text{si } k = n\,; \\ 0, & \text{si } k \neq n. \end{cases}$$

Montrer que pour tout x dans $\ell^2(\mathbb{N})$, la suite $(A_n x)$ converge vers 0, que $\|A_n\| = 1$, pour tout $n \in \mathbb{N}$ et que la suite (A_n) ne converge pas en norme. Y a-t-il contradiction avec le résultat de l'exercice précédent ?

Solution : Pour tout élément x de $\ell^2(\mathbb{N})$, on a

$$x = \sum_{n=0}^{\infty} \langle x, e_n \rangle e_n \quad \text{et} \quad \|x\|^2 = \sum_{n=0}^{\infty} |\langle x, e_n \rangle|^2$$

On en déduit en particulier que la suite $(\langle x, e_n \rangle)$ tend vers 0. Par définition de l'opérateur A_n, on a

$$\forall x \in \ell^2(\mathbb{N}), \quad A_n x = \langle x, e_n \rangle e_n$$

On en déduit que la suite $(A_n x)$ converge vers 0, elle y converge donc simplement ; on en déduit aussi que

$$\forall x \in \ell^2(\mathbb{N}), \quad \|A_n x\| \leq \|x\|$$

Cela montre que la norme de A_n est plus petite que 1. Comme pour $x = e_n$, on a $A_n e_n = e_n$, il en résulte que, pour tout entier n, $\|A_n\| = 1$. Cependant, si la suite (A_n) convergeait vers un opérateur A, on aurait d'un côté $\|A\| = 1$ et d'un autre côté

$$\forall x, \quad Ax = \lim_{n \to +\infty} A_n x = 0$$

Ce qui est impossible. □

7. Soit E un espace de Hilbert sur $\mathbb{C}$ et soit A un élément de $\mathcal{L}(E)$.

(a) Montrer que

$$\|A\| = \sup_{\|x\|, \|y\| \leq 1} |\langle Ax, y \rangle|$$

(b) Montrer que si deux éléments A et B de $\mathcal{L}(E)$ vérifient la relation $\langle Ax, y \rangle = \langle Bx, y \rangle$ quels que soient x et y dans E, alors ils sont égaux.

(c) Prouver l'identité de polarisation suivante :

$$\langle Ax, y \rangle = \frac{1}{4} \sum_{n=0}^{3} i^n \langle A(x + i^n y), x + i^n y \rangle$$

(d) En déduire que $A = B$ si, et seulement si, $\langle Ax, x \rangle = \langle Bx, x \rangle$, pour tout x dans E.

(e) Ce dernier résultat reste-t-il encore vrai pour les espaces de Hilbert réels ?

8. Sur l'espace de Hilbert $L^2([0, \infty[, dx)$, on définit l'opérateur A par

$$Af(x) = \frac{1}{x} \int_0^x f(t)\, dt$$

Montrer que A est borné.

Solution : Pour f définie sur $(0,\infty)$, on associe la fonction Af définie sur $\mathbb{R}$ par $Af(s) = f(e^s)e^{s/2}$, on a donc $f(t) = Af(\log t)t^{-1/2}$, $t > 0$. On vérifie que A est une bijection de $L^2(]0,\infty[,dx)$ sur $L^2(\mathbb{R},dx)$ isométrique c'est-à-dire que $\|f\|_{L^2(0,\infty)} = \|Af\|_{L^2(\mathbb{R})}$:

$$\int_0^\infty |f(t)|^2\,dt = \int_{\mathbb{R}} |Af(s)|^2\,ds$$

On vérifie ensuite que $A(Tf) = Af * E$ où $E(s) = e^{-s/2}\chi(s)$ et χ est la fonction caractéristique de $(0,\infty)$. Cela veut dire que l'opérateur T se transforme par A en l'opérateur de convolution par E ; celui-ci est borné et a pour norme 2, car $\|E\|_1 = 2$, il en résulte que T est borné et sa norme vaut 2. □

3.2 Exemples d'opérateurs linéaires continus

Voyons maintenant quelques exemples d'opérateurs linéaires continus d'un espace de Hilbert dans lui-même.

Opérateurs de multiplication et de translation

Exemple 3.2.1. - Prenons $E = \ell^2(\mathbb{N})$ et soit (λ_n), $n \in \mathbb{N}$, une suite bornée de nombres complexes. Pour une suite $\mathbf{x} = (x_n)$ dans $\ell^2(\mathbb{N})$, on pose $T\mathbf{x} = (\lambda_n x_n)$. On vérifie immédiatement que

$$\|T\mathbf{x}\|^2 = \sum_{n=0}^{\infty} |\lambda_n x_n|^2 \leq \sup_n |\lambda_n|^2\ \|\mathbf{x}\|^2$$

L'opérateur T est donc continu sur $\ell^2(\mathbb{N})$ et $\|T\| \leq \sup_n |\lambda_n|$. En fait (voir exercice 1) on a l'égalité $\|T\| = \sup_n |\lambda_n|$. T est appelé *opérateur de multiplication* par la suite $\lambda = (\lambda_n)$.

Exemple 3.2.2. - Prenons $E = L^2(X,\Omega,\mu)$, où (X,Ω,μ) est un espace mesuré σ-fini et soit ϕ une fonction mesurable sur X à valeurs complexes. On suppose que ϕ est essentiellement bornée, c'est-à-dire qu'il existe une constante $c \geq 0$ telle que $|\phi(x)| \leq c$ pour presque tout x de X. La plus petite de ces constantes c, notée $\|\phi\|_\infty$, est appelée la borne supérieure essentielle de ϕ. Pour f dans $L^2(X,\Omega,\mu)$, on pose $Tf = \phi f$. On définit ainsi l'*opérateur de multiplication par* ϕ et on vérifie que sa norme vaut $\|\phi\|_\infty$.

Par exemple sur $L^2([0,1], dx)$ l'application qui à f fait correspondre xf est un opérateur linéaire continu et sa norme est égale à 1.
Remarquons que l'égalité $\|T\| = \|\phi\|_\infty$ tombe en défaut si l'espace X n'est pas σ-fini, comme le montre l'exemple suivant : $X = [0,1]$ muni de la tribu borélienne Ω et de la mesure

$$\Delta \in \Omega \mapsto \mu(\Delta) = \begin{cases} m(\Delta) & \text{si } 0 \notin \Delta\,; \\ +\infty & \text{si } 0 \in \Delta. \end{cases}$$

où m est la mesure de Lebesgue. La mesure μ n'est pas σ-finie, puisque $\mu(\{0\}) = +\infty$. Prenons ϕ la fonction caractéristique de 0; elle est essentiellement bornée et $\|\phi\|_\infty = 1$. Pour tout $f \in L^2(\mu)$, l'inégalité

$$\int |f|^2 d\mu \geq |f(0)|\mu(\{0\})$$

implique nécessairement $f(0) = 0$ et donc $\phi f = 0$. L'opérateur de multiplication par ϕ est donc nul alors que ϕ n'est pas nulle presque-partout.

Exemple 3.2.3. - Pour tout réel a, on désigne par T_a l'opérateur qui, à une fonction f de $L^2(\mathbb{R})$, associe la fonction $T_a f$ définie presque-partout par

$$T_a f(x) = f(x-a)$$

C'est un opérateur linéaire continu sur $L^2(\mathbb{R})$ et sa norme est égale à 1 ; il est appelé *opérateur de translation*.

Exemple 3.2.4. - Pour un entier relatif p, on désigne par T_p l'opérateur qui, à une suite $\mathbf{x} = (x_n)$ de $\ell^2(\mathbb{Z})$, fait correspondre la suite

$$T_p \mathbf{x} = (x_{n-p})$$

On a là un opérateur linéaire continu sur $\ell^2(\mathbb{Z})$ et de norme 1. Il est de même appelé *opérateur de translation*.

Opérateur intégral de Fredholm[1]

Considérons un intervalle fermé borné $[a,b]$ et une fonction k continue sur $[a,b] \times [a,b]$ à valeurs dans $\mathbb{C}$. A une fonction f de $L^2([a,b], dx)$, on fait correspondre la fonction Kf définie sur l'intervalle $[a,b]$ par

$$Kf(x) = \int_a^b k(x,y) f(y)\, dy$$

[1]Ivar Fredholm (1866-1927), est un mathématicien suédois dont le nom reste attaché à la théorie des équations intégrales. Ses travaux sont à l'origine des travaux de Hilbert, sur le même sujet, ce qui allait conduire celui-ci à la notion fondamentale d'espace de Hilbert.

Nous allons montrer que K est un opérateur linéaire continu de $L^2([a,b],dx)$ dans lui-même. En effet, une application de l'inégalité de Cauchy-Schwarz permet d'écrire

$$|Kf(x)|^2 \leq \left(\int_a^b |k(x,y)|^2 dy\right) \|f\|^2$$

Par intégration, on en déduit que $\|Kf\| \leq M(b-a)\|f\|$, où M désigne le maximum de la fonction k sur $[a,b] \times [a,b]$. Ainsi, on a montré que l'opérateur K est continu et que $\|K\| \leq M(b-a)$. K est appelé *opérateur intégral de Fredholm* et la fonction k est appelée le *noyau* de l'opérateur K. Il faut préciser que, pour f dans $L^2([a,b],dx)$, la fonction Kf est continue sur $[a,b]$. En effet, la fonction $(x,y) \mapsto k(x,y)f(y)$ est continue en x et pour presque tout y elle peut être majorée indépendamment de x par une fonction intégrable sur $[a,b]$, à savoir $M|f|$; le théorème de convergence dominée s'applique donc et permet de conclure. Il arrive souvent que l'intervalle dans lequel on travaille ne soit pas borné ou que le noyau k ne soit pas continu. Le théorème suivant permet d'envisager des exemples de cette situation.

Théorème 3.2.5. *Soit (X,Ω,μ) un espace mesuré σ-fini et soit k une fonction mesurable, définie sur $X \times X$ à valeurs complexes. On suppose qu'il existe deux constantes c_1 et c_2 telles que*

$$\int_X |k(x,y)|d\mu(y) \leq c_1, \qquad \mu - \textit{presque partout}$$

et

$$\int_X |k(x,y)|d\mu(x) \leq c_2, \qquad \mu - \textit{presque partout}$$

Alors pour toute f dans $L^2(X,\Omega,\mu)$, la fonction

$$Kf(x) = \int_X k(x,y)f(y)d\mu(y)$$

est définie presque-partout, appartient à $L^2(X,\Omega,\mu)$ et l'opérateur K qui à f fait correspondre Kf est continu dans $L^2(X,\Omega,\mu)$ et sa norme vérifie $\|K\| \leq (c_1c_2)^{\frac{1}{2}}$.

Démonstration. On doit démontrer que Kf appartient à $L^2(\mu)$, mais ceci découlera de la démonstration du fait que K est borné. Si $f \in L^2(\mu)$, l'inégalité de Hölder montre que, pour presque tout x,

$$\begin{aligned}|Kf(x)| &\leq \int |k(x,y)|\,|f(y)|d\mu(y) \\ &\leq \Big[\int |k(x,y)|d\mu(y)\Big]^{\frac{1}{2}} \Big[\int |k(x,y)|\,|f(y)|^2 d\mu(y)\Big]^{\frac{1}{2}} \\ &\leq \sqrt{c_1}\Big[\int |k(x,y)|\,|f(y)|^2 d\mu(y)\Big]^{\frac{1}{2}}\end{aligned}$$

donc

$$\begin{aligned}\int |Kf(x)|^2 d\mu(x) &\leq c_1 \int\int |k(x,y)|\,|f(y)|^2 d\mu(y) d\mu(x)\\ &\leq c_1 \int |f(y)|^2 \int |k(x,y)| d\mu(x) d\mu(y)\\ &\leq c_1 c_2 \|f\|^2.\end{aligned}$$

Cela montre que la fonction Kf est définie μ-presque partout, appartient à $L^2(X, \Omega, \mu)$ et que $\|Kf\|^2 \leq c_1 c_2 \|f\|^2$. □

L'opérateur décrit par ce théorème est aussi appelé *opérateur intégral* de *noyau k*. Le corollaire suivant en fournit plusieurs exemples intéréssants.

Corollaire 3.2.6. *Soit φ une fonction mesurable sur $\mathbb{R}^n$ et soit la fonction k définie sur $\mathbb{R}^n \times \mathbb{R}^n$ par $k(x,y) = \varphi(x-y)$. On suppose que φ est intégrable sur $\mathbb{R}^n$ muni de la mesure de Lebesgue, alors l'application*

$$f \mapsto \varphi * f(x) = \int_{\mathbb{R}^n} \varphi(x-y) f(y)\, dy$$

définit un opérateur linéaire K_φ, borné de $L^2(\mathbb{R}^n)$ dans lui-même. De plus, on a $\|K_\varphi\| \leq \|\varphi\|_{L^1(\mathbb{R}^n)}$.

Les opérateurs de ce type, appelés *opérateurs de convolution*, interviennent dans beaucoup de questions d'analyse. En voici quelques exemples bien connus.

Exemple 3.2.7. - *Opérateur de Poisson*[2] *sur* $\mathbb{R}$: C'est l'opérateur de convolution associé à la fonction

$$\varphi(x) = \frac{1}{\pi} \frac{1}{(1+x^2)}$$

Le coefficient $1/\pi$ a été ajusté pour que $\|\varphi\|_1 = 1$. L'opérateur de Poisson, que nous noterons pour la circonstance par P, est alors défini par

$$Pf(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(y)}{1+(x-y)^2}\, dy, \quad \text{où } f \in L^2(\mathbb{R}, dx)$$

[2]Siméon-Denis Poisson (1781–1840), mathématicien français, abondonna ses études de médecine auxquelles ses parents voulaient l'orienter, pour aller étudier les mathématiques à l'Ecole polytechnique, où il fut élève de P. Laplace et J. Lagrange qui devinrent l'un et l'autre ses amis. Ses travaux portent sur les intégrales définies, la théorie électromagnétique et le calcul des probabilités. Le nom de Siméon Denis Poisson est attaché à de nombreuses notions mathématiques et physiques (intégrale et équation de Poisson en théorie du potentiel, crochets de Poisson dans la théorie des équations différentielles, rapport de Poisson en élasticité et constante de Poisson en électricité).

C'est un opérateur linéaire borné de $L^2(\mathbb{R}, dx)$ dans lui-même et de norme égale à 1. On sait l'importance de cet opérateur dans la théorie des fonctions harmoniques ; plus précisément, posons pour $t > 0$

$$\varphi_t(x) = \frac{1}{t}\varphi(x/t) = \frac{1}{\pi}\frac{t}{t^2 + x^2}$$

La fonction φ_t est d'intégrale 1 et l'opérateur de convolution P_t qui lui est associé :

$$P_t f(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{tf(y)}{t^2 + (x-y)^2}\,dy$$

est continu de $L^2(\mathbb{R}, dx)$ dans lui-même. La fonction $(x, t) \mapsto \varphi_t(x)$, définie dans le demi-plan supérieur $\{ (x, t); x \in \mathbb{R}, t > 0 \}$ est la partie imaginaire de $-1/(\pi z)$, où $z = x + it$, et est donc une fonction harmonique, c'est-à-dire qu'elle vérifie

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial t^2}\right)\varphi_t(x) = 0, \quad \text{pour} \quad x \in \mathbb{R}, \quad \text{et} \quad t > 0$$

Par dérivation sous le signe d'intégration, qui se justifie facilement à l'aide du théorème de convergence dominée, on voit que la fonction $(x, t) \mapsto P_t f(x)$ est harmonique dans le demi-plan supérieur, de plus au sens de la norme de $L^2(\mathbb{R}, dx)$, on a

$$\lim_{t \to 0^+} \|P_t f - f\| = 0$$

Ceci est un résultat général qui découle du fait que $\{ \varphi_t, t > 0 \}$ est une approximation de l'identité (voir exercice 1). De là , on peut voir que la convergence a lieu en tout point où f est continue (et dans $L^q(\mathbb{R})$, $1 \leq q < \infty$, si f est dans cet espace). Ainsi, $P_t f$ résoud le problème de Dirichlet dans le demi-plan supérieur :

$$\begin{cases} \left(\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial t^2}\right) P_t f(x) = 0, & \text{pour } t > 0\,; \\ \lim_{t \to 0} P_t f = f, & \text{au sens ci-dessus.} \end{cases}$$

Exemple 3.2.8. - *Opérateur de Poisson du disque unité* : Pour $0 \leq r < 1$, on pose

$$p_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}$$

C'est une fonction 2π-périodique, continue, paire, positive et on a

$$p_r(\theta) = \sum_{n \in \mathbb{Z}} r^{|n|} e^{in\theta}$$

En effet, le second membre s'écrit

$$\sum_{n\in\mathbb{Z}} r^{|n|}e^{in\theta} = 1 + \sum_{n=1}^{\infty}\left(r^n e^{in\theta} + r^n e^{-in\theta}\right)$$

et il suffit alors de se rappeler la formule donnant la somme d'une série géométrique. Il est clair que

$$\frac{1}{2\pi}\int_0^{2\pi} p_r(\theta)\, d\theta = 1$$

L'opérateur de convolution $\mathcal{P}_r$ qui est associé à la fonction p_r :

$$\mathcal{P}_r f(\theta) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta')p_r(\theta-\theta')\, d\theta'$$

est un opérateur linéaire continu de $L^2([0,2\pi], dx/2\pi)$ dans lui-même et sa norme vaut 1. La formule de Parseval (voir le théorème 2.1, section 2, chapitre II) permet d'exprimer $\mathcal{P}_r f$, pour f dans $L^2([0,2\pi], dx/2\pi)$, à l'aide des coefficients de Fourier de p_r et ceux de f :

$$\begin{aligned}
\mathcal{P}_r f(\theta) &= \sum_{n\in\mathbb{Z}} \hat{f}(n) r^{|n|} e^{in\theta} \\
&= \sum_{n=0}^{\infty} \hat{f}(n) z^n + \sum_{n=1}^{\infty} \hat{f}(-n)\bar{z}^n \\
&= v(z) + w(z), \qquad \text{avec} \quad z = re^{i\theta}
\end{aligned}$$

Rappelons que f étant de carré intégrable, ses coefficients de Fourier $(\hat{f}(n))$ sont bornés (en fait $\hat{f}(n)$ tend vers 0 quand $|n|$ tend vers l'infini) et donc les deux séries ci-dessus ont un rayon de convergence au moins égal à 1. Il est clair que, dans le disque unité, la fonction $v(z)$ est holomorphe et la fonction $w(z)$ est antiholomorphe. En d'autres termes, les fonctions v et w sont indéfiniment différentiables dans le disque unité $D = \{\, z \in \mathbb{C} \mid |z| < 1 \,\}$ et

$$\frac{\partial v}{\partial \bar{z}} = 0, \quad \frac{\partial w}{\partial z} = 0, \quad \text{sur } D$$

où l'on a posé

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right), \qquad \frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)$$

Notons que

$$\frac{\partial}{\partial z}\frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial \bar{z}}\frac{\partial}{\partial z} = \frac{1}{4}\Delta$$

où Δ est l'opérateur de Laplace sur $\mathbb{R}^2$. Puisque v et w sont dans $C^\infty(D)$, nous avons $\Delta v = 0$ et $\Delta w = 0$, et donc

$$\Delta \mathcal{P}_r f = 0$$

De plus, si l'on désigne par $\| \ \|$ la norme de $L^2([0, 2\pi], dx/2\pi)$, la formule de Parseval donne

$$\|f - \mathcal{P}_r f\|^2 = \sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2 (1 - r^{|n|})^2$$

et par suite

$$\lim_{r \to 1^-} \|f - \mathcal{P}_r f\| = 0$$

Ainsi, $P_r f$ est la solution du problème de Dirichlet dans le disque unité en ce sens que, pour toute f dans $L^2([0, 2\pi], dx/2\pi)$, la fonction $P_r f$, $r < 1$, est une fonction harmonique à l'intérieur du disque unité et converge vers f sur la frontière de D, lorsque r tend vers 1^- (la convergence étant entendue au sens de la norme de l'espace $L^2([0, 2\pi], dx/2\pi)$).

Exemple 3.2.9. - *Opérateur de Gauss*[3] Considérons la fonction de Gauss :

$$g(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2), \quad \text{pour} \quad x \in \mathbb{R}$$

La fonction g est intégrable sur $\mathbb{R}$ et son intégrale vaut 1. L'opérateur de convolution G, qui lui est associé :

$$Gf(x) = \int_{\mathbb{R}} f(y) g(x - y)\, dy = \int_{\mathbb{R}} f(x - y) g(y)\, dy$$

est borné de $L^2(\mathbb{R}, dx)$ dans lui-même et sa norme est majorée par 1. Posons, pour $t > 0$,

$$g_t(x) = \frac{1}{\sqrt{t}} g\left(x/\sqrt{t}\right)$$

La fonction g_t est d'intégrale égale à 1 et l'opérateur de convolution associé, G_t est borné sur $L^2(\mathbb{R}, dx)$; il est appelé opérateur de Gauss et joue un rôle important dans la résolution des équations d'évolution classiques telle que l'équation de la chaleur. En effet, on vérifie aisément que

$$\frac{\partial g_t}{\partial t} = \frac{1}{2} \frac{\partial^2 g_t}{\partial x^2}$$

[3]L'œuvre du mathématicien allemand Carl Friedrich Gauss (1777-1855) est un monument d'une ampleur et d'une richesse sans égale : non seulement il y a Gauss mathématicien, mais il y a le calculateur, le géodésien, l'astronome sans oublier qu'il a pratiquement consacré les vingt dernières années de sa vie à l'étude du magnétisme. On estime que c'est presque de trois quarts de siècle qu'il a devancé son temps et ainsi illuminé l'avenir comme nul autre ne l'a fait. Son génie inspirait à ses contemporains une vénération un peu craintive, et nul n'aurait osé lui contester le titre de "Prince des mathématiciens" dont on le désignait communément.

On en déduit que, pour toute f dans $L^2(\mathbb{R}, dx)$, la fonction qui à (x, t) associe $G_t f(x)$ est deux fois continûment dérivable sur $\mathbb{R}\times]0, \infty[$ et satisfait

$$\begin{cases} \dfrac{\partial G_t}{\partial t} = \dfrac{1}{2}\dfrac{\partial^2 G_t}{\partial x^2} \\ \lim\limits_{t\to 0^+} \|f - G_t f\| = 0 \end{cases}$$

c'est-à-dire que l'opérateur de Gauss permet de résoudre l'équation de la chaleur sur $\mathbb{R}\times]0, \infty[$.

Opérateur intégral de Volterra[4]

Soit $[a, b]$ un intervalle fermé borné de $\mathbb{R}$ et soit k une fonction continue sur $[a, b] \times [a, b]$ à valeurs complexes. Pour toute fonction f dans l'espace $L^2([a, b], dx)$, on pose

$$Vf(x) = \int_a^x k(x, y) f(y)\, dy$$

On définit ainsi un opérateur linéaire borné de $L^2([a, b], dx)$ dans lui-même et on vérifie que

$$\|V\| \leq M\frac{b-a}{\sqrt{2}} \quad \text{où} \quad M = \sup_{[a,b]\times[a,b]} |k(x, y)|$$

Cet opérateur, appelé opérateur intégral de Volterra, diffère de l'opérateur intégral de Fredholm par le fait que la borne supérieure de l'intégrale qui le définit est la variable x, au lieu de l'extrémité b de l'intervalle. Ceci a pour effet d'avoir une meilleure majoration de la norme de V et par suite de V^n, avec $n \in \mathbb{N}$. On peut anticiper et voir l'exemple 3.12 du paragraphe qui suit.

EXERCICES

1. Soit $\lambda = (\lambda_n)$ une suite bornée de nombres complexes et soit T_λ l'opérateur de multiplication dans $\ell^2(\mathbb{N})$, par la suite λ. Montrer que $\|T_\lambda\| = \sup_n |\lambda_n|$.

Solution : Par définition, pour $x = (x_n) \in \ell^2(\mathbb{N})$, $T_\lambda x = (\lambda_n x_n)$ et

$$\|T_\lambda x\|^2 = \sum_{n\in\mathbb{N}} |\lambda_n x_n|^2 \leq \sup_{n\in\mathbb{N}} |\lambda_n|^2 \|x\|^2$$

[4]Vito Volterra (1860-1940), est un mathématicien italien dont les travaux portent sur l'Analyse Mathématique et ses applications à la Mécanique Physique et la Biologie. Il consacra les dernières années de sa vie à l'étude de l'Hérédité et la Théorie de la lutte pour la vie.

On en déduit que $\|T_\lambda\| \leq \sup_{n\in\mathbb{N}} |\lambda_n|$. D'autre part, pour tout élément e_k de la base canonique de $\ell^2(\mathbb{N})$, on a $T_\lambda e_k = \lambda_k e_k$ et par suite

$$\forall k \in \mathbb{N}, \quad \|T_\lambda\| \geq \frac{\|T_\lambda e_k\|}{\|e_k\|} = |\lambda_k|$$

On en déduit que $\sup_{k\in\mathbb{N}} |\lambda_k| \leq \|T_\lambda\|$, et par suite l'égalité voulue. □

2. *Approximation de l'identité* : Soit φ une fonction positive, intégrable sur $\mathbb{R}^n$ et d'intégrale égale à 1. Pour $t > 0$, on considère la fonction définie par $\varphi_t(x) = t^{-n}\varphi(x/t)$

(i) Montrer que φ_t est dans $L^1(\mathbb{R}^n)$ et que $\|\varphi_t\|_1 = 1$.

(ii) Montrer que pour tout $\delta > 0$, $\lim_{t\to 0} \int_{\|x\|>\delta} \varphi_t(x)\,dx = 0$.

(iii) En déduire que si f est dans $L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, alors $\varphi_t * f$ appartient à $L^p(\mathbb{R}^n, dx)$ et que

$$\lim_{t\to 0^+} \|f - \varphi_t * f\|_p = 0$$

(iv) Montrer que si f est une fonction uniformément continue et bornée sur $\mathbb{R}^n$, alors $\varphi_t * f$ converge uniformément sur $\mathbb{R}^n$ vers f lorsque t tend vers 0.

Solution : L'assertion (i) résulte immédiatement du changement de variables $y = x/\epsilon$. Pour l'assertion (ii), soit $\delta > 0$ et posons encore $y = x/\epsilon$. Alors

$$\int_{\|x\|>\delta} \varphi_t(x)\,dx = \epsilon^{-n} \int_{\|y\|>\delta} \varphi(x/\epsilon)\,dx = \int_{\|y\|>\delta/\epsilon} \varphi(y)\,dy$$

Puisque φ est intégrable et δ/ϵ tend vers l'infini lorsque ϵ tend vers 0, la dernière intégrale tend vers 0 avec ϵ. Pour la partie (iii), on remarque d'après (i) que $f(x) = \int_{\mathbb{R}^n} \varphi_t(x)\,dx$ et par suite

$$\begin{aligned} |\varphi_t * f| &= \left| \int_{\mathbb{R}^n} [f(x-y) - f(x)]\varphi_t(y)\,dy \right| \\ &\leq \int_{\mathbb{R}^n} |f(x-y) - f(x)| \varphi_t(y)^{1/p} \varphi_t(y)^{1/q}\,dy \end{aligned}$$

où $1/p + 1/q = 1$. En appliquant l'inégalité de Hölder et en intégrant par rapport à x, on voit que $\|f * \varphi_t(x) - f(x)\|_p^p$ est majorée par

$$\begin{aligned} &\int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x-y) - f(x)|^p \varphi_t(y)\,dy \right] \left[\int_{\mathbb{R}^n} \varphi_t(y)\,dy \right]^{p/q} dx \\ &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x-y) - f(x)|^p \varphi_t(y)\,dy \right] dx \end{aligned}$$

En changeant l'ordre des intégrations dans la dernière intégrale (ce qui est justifié car les fonctions sont positives), on obtient

$$\|\varphi_t * f - f\|_p^p \leq \int_{\mathbb{R}^n} \varphi_t(y)\psi(y)\, dy$$

où $\psi(y) = \|T_y f - f\|_p^p$, T_y étant l'opérateur de translation défini au début de ce paragraphe. Pour $\delta > 0$, on peut écrire

$$\int_{\mathbb{R}^n} \varphi_t(y)\psi(y)\, dy = \int_{\|y\|<\delta} + \int_{\|y\|\geq\delta} \varphi_t(y)\psi(y)\, dy = A_{t,\delta} + B_{t,\delta}$$

Etant donné $\eta > 0$, on peut choisir δ de façon que $\psi(y)$ soit strictement inférieur à η dès que $\|y\| < \delta$; il en résulte que pour tout $t > 0$, $A_{t,\delta} \leq \eta$. De plus, en utilisant l'inégalité de Minkowski, on peut voir que $\|\psi\|_\infty$ est majoré par $2^p\|f\|_p^p$ et par suite $B_{t,\delta}$ est, à une constante multiplicative près, majorée par $\int_{\|y\|\geq\delta} \varphi_t(y)\, dy$. On conclut grâce à (ii). L'assertion (iv) se démontre de façon tout à fait analogue. □

3. En utilisant les résultats de l'exercice précédent, montrer que l'espace $C_c^\infty(\mathbb{R}^n)$ des fonctions indéfiniment dérivables et à support compact dans $\mathbb{R}^n$ est dense dans $L^p(\mathbb{R}^n, dx)$ pour tout p, $1 \leq p < \infty$.

4. *Lemme de Schur* : Soit k une fonction définie et positive sur $\mathbb{R}_+ \times \mathbb{R}_+$. On suppose que pour tout $\lambda > 0$, $k(\lambda s, \lambda t) = \lambda^{-1}k(s,t)$ et que pour un certain p, $1 \leq p \leq \infty$,

$$\int_0^\infty k(1,t)t^{-\frac{1}{p}}\, dt = \gamma < \infty$$

Par exemple $k(s,t) = 1/(s+t)$ possède ces propriétés. Montrer que $\|Kf\|_p \leq \gamma\|f\|_p$, où K est donné par

$$Kf(s) = \int_0^\infty f(t)k(s,t)\, dt, \quad (s > 0)$$

Solution : On remarque que

$$Kf(s) = s^{-1}\int_0^\infty f(t)k(1,t/s)\, dt = \int_0^\infty f(st)k(1,t)\, dt$$

En utilisant la version intégrale de l'inégalité de Minkowski (voir exercice 15, section 1, chapitre I), on peut écrire

$$\|Kf\|_p \leq \int_0^\infty \left[\int_0^\infty |f(st)k(1,t)|^p ds\right]^{\frac{1}{p}} dt$$

Un changement de variables évident montre que le second membre est égal à

$$\int_0^\infty k(1,t)t^{-\frac{1}{p}}\left(\int_0^\infty |f(u)|^p\, du\right)^{\frac{1}{p}} dt = \gamma\|f\|_p$$

L'inégalitée cherchée s'en déduit. □

5. (*Extension, due à Schur, du théorème 2.5*) : Soient (X, μ) et (Y, ν) deux espaces mesurés σ-finis et soit $(x, y) \mapsto k(x, y)$ une fonction mesurable sur $(X \times Y, \mu \times \nu)$. On suppose qu'il existe des fonctions mesurables $p(x)$ et $q(y)$ positives presque partout sur X et Y respectivement et telles que

$$\int_X |k(x,y)| p(x)\, d\mu(x) \leq C_1 q(y), \quad \int_Y |k(x,y)| q(y)\, d\nu(y) \leq C_2 p(x)$$

(i) Montrer que l'opérateur intégral de noyau k, défini par

$$Kf(x) = \int_Y k(x,y) f(y)\, d\nu(y)$$

est borné de $L^2(Y, \nu)$ dans $L^2(X, \mu)$ et que $\|K\|^2 \leq C_1 C_2$.
(ii) Montrer que dans le cas où $X = \mathbb{R}_+ = [0, \infty[$ et $Y = \mathbb{R}$, l'opérateur K est borné de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R}_+)$ pourvu que le noyau k vérifie l'inégalité $|k(x, y)| \leq C\,(x^2 + y^2)^{-\frac{1}{2}}$. On construira $p(x)$ et $q(y)$ de façon à pouvoir utiliser (i).

3.3 Propriétés spectrales des opérateurs continus

Soit E un espace de Hilbert, $\mathcal{L}(E)$ est alors une algèbre de Banach. On aimerait définir $f(A)$ lorsque f est une fonction d'une variable complexe et A un élément de $\mathcal{L}(E)$. Prenons d'abord f un polynôme à coefficients complexes, $f(x) = \sum_0^n \alpha_k x^k$. Il n'y a aucune difficulté à définir $f(A)$ par

$$f(A) = \sum_0^n \alpha_k A^k$$

en convenant que $A^0 = I$ (l'opérateur identité) et $f(A)$ appartient à $\mathcal{L}(E)$. On vérifie également que si p_1 et p_2 sont deux polynômes à coefficients complexes

$$(\lambda_1 p_1 + \lambda_2 p_2)(A) = \lambda_1 p_1(A) + \lambda_2 p_2(A), \quad p_1 p_2(A) = p_1(A) p_2(A)$$

ce qui exprime le fait que l'application $p \mapsto p(A)$ est un homomorphisme de l'algèbre des polynômes dans l'algèbre $\mathcal{L}(E)$.

Jusqu'ici nous n'avons pas fait usage de topologie. Prenons maintenant pour f la somme d'une série entière $f(x) = \sum_0^\infty \alpha_j x^j$, de rayon de convergence $R > 0$; la relation $\|A^n\| \leq \|A\|^n$, montre que la série $\sum_0^\infty \alpha_j A^j$ est

normalement convergente pour $\|A\| < R$. Comme $\mathcal{L}(E)$ est complet, elle est convergente et sa somme définit un élément de $\mathcal{L}(E)$ qu'on note $f(A)$. On peut par exemple considérer la série

$$\sum_{n=0}^{\infty} \frac{A^n}{n!},$$

dont le rayon de convergence est $R = +\infty$. Sa somme est un opérateur borné noté e^A, et on vérifie facilement que si A et B sont deux éléments de $\mathcal{L}(E)$ qui commutent ($AB = BA$), alors $e^A e^B = e^{(A+B)}$. En particulier si l'on pose, pour $t \in \mathbb{C}$, $f(t) = e^{tA}$, on aura

$$f(t+s) = f(t)f(s)$$

et on a là un groupe à un paramètre d'éléments de $\mathcal{L}(E)$.
Un autre exemple particulièrement important est donné par le théorème suivant

Théorème 3.3.1. *Si A est un élément de $\mathcal{L}(E)$ tel que $\|A\| < 1$, alors la série $\sum_{n=0}^{\infty} A^n$ est convergente dans $\mathcal{L}(E)$ et on a*

$$(I - A)\Big(\sum_{n=0}^{\infty} A^n\Big) = \Big(\sum_{n=0}^{\infty} A^n\Big)(I - A) = I$$

Démonstration. Si $\|A\| < 1$ la série $\sum A^n$ est normalement convergente dans $\mathcal{L}(E)$. Soit B sa somme ; la relation

$$(I - A)(I + A + A^2 + \cdots + A^n) = I - A^{n+1}$$

montre, par passage à la limite quand n tend vers l'infini, que B vérifie $(I - A)B = B(I - A) = I$. □

Ce théorème traduit le fait que si $\|A\| < 1$, l'opérateur $I - A$ est inversible et son inverse est la somme de la série $\sum_0^{\infty} A^n$. Celle-ci est souvent appelée série de Neumann[5], elle a été introduite dans le but de résoudre certaines équations intégrales dont nous verrons, plus loin, quelques exemples.

Rappelons qu'un élément A de $\mathcal{L}(E)$ est dit *inversible* s'il existe un élément B (nécessairement dans $\mathcal{L}(E)$, voir remarque ci-dessous) vérifiant

$$AB = I \quad \text{et} \quad BA = I$$

[5] John Von NEUMANN (1903-1957), mathématicien d'origine hongroise, a créé des théories nouvelles pour formaliser des problèmes qui échappaient à l'analyse classique. Il a ainsi rendu rigoureux les développements du formalisme de la Mécanique Quantique de Dirac et Schrödinger. Il concocta en 1943, avec l'économiste autrichien Oskar Morgenstern, une célèbre théorie économique : la théorie des jeux. Les ordinateurs d'aujourd'hui fonctionnent toujours selon les principes de base qu'il définit en juin 1945.

L'opérateur B est alors unique ; il est désigné par $B = A^{-1}$ et appelé l'inverse de A. Notons qu'une seule des égalités précédentes ne suffit pas pour affirmer que A est inversible. En effet la première implique que l'opérateur A est surjectif alors que la deuxième implique qu'il est injectif. Si E est de dimension finie, il y a effectivement équivalence entre la première et la deuxième relation, mais si E n'est pas de dimension finie, cette équivalence tombe en défaut.

Remarque 3.3.2. - C'est un fait remarquable que si $A \in \mathcal{L}(E)$ est bijectif, alors son inverse A^{-1} est un opérateur continu, c'est-à-dire appartient à $\mathcal{L}(E)$. Ceci est une conséquence du théorème de l'application ouverte (voir l'annexe) qui assure que si E et F sont deux espaces de Banach et $\phi : E \mapsto F$ est une application linéaire continue et surjective, alors elle est ouverte, c'est-à-dire qu'elle envoie tout ouvert de E sur un ouvert de F. Il en résulte que l'image réciproque par A^{-1} de tout ouvert $\mathcal{O} \subset E$, c'est-à-dire $A(\mathcal{O})$, est un ouvert ce qui traduit précisément la continuité de A^{-1}.

Lemme 3.3.3. *L'ensemble des éléments inversibles de $\mathcal{L}(E)$ est un ouvert dans $\mathcal{L}(E)$ et l'application $A \mapsto A^{-1}$ est continue.*

Démonstration. Soit A un élément de $\mathcal{L}(E)$ inversible. Pour tout l'élément B de $\mathcal{L}(E)$ vérifiant $\|B\| < \|A^{-1}\|^{-1}$, l'opérateur $I + A^{-1}B$ est inversible (car $\|A^{-1}B\| < 1$) et il en sera de même de l'opérateur $A + B = A(I + A^{-1}B)$. Cela montre que les opérateurs inversibles forment un ouvert de $\mathcal{L}(E)$. Par ailleurs,

$$(A+B)^{-1} - A^{-1} = (I + A^{-1}B)^{-1}A^{-1} - A^{-1}$$
$$= \sum_{n=1}^{\infty}(-1)^n (A^{-1}B)^n A^{-1}$$

et par suite

$$\|(A+B)^{-1} - A^{-1}\| \leq \sum_{1}^{\infty} \|A^{-1}\|^{n+1}\|B\|^n < \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|B\|}\|B\|$$

Il en résulte que $\|(A+B)^{-1} - A^{-1}\|$ tend vers 0 avec $\|B\|$. □

Définition 3.3.4. *Soit A un élément de $\mathcal{L}(E)$. On appelle ensemble résolvant de, et on note $\rho(A)$, l'ensemble des nombres complexes λ tels que $A - \lambda I$ soit inversible. On appelle spectre de A, et on note $\sigma(A)$, le complémentaire de $\rho(A)$ dans $\mathbb{C}$.*

L'application R_A qui à $\lambda \in \rho(A)$ fait correspondre $(A - \lambda I)^{-1}$, est appelée la résolvante de l'opérateur A. Elle possède la propriété intéressante suivante, qui permet de donner une première description du spectre d'un opérateur de $\mathcal{L}(E)$.

Théorème 3.3.5. *Soit E un espace de Hilbert et $A \in \mathcal{L}(E)$.*

(i) *La résolvante R_A est une application holomorphe de $\rho(A)$ dans $\mathcal{L}(E)$ et vérifie l'équation de la résolvante*

$$R_A(\lambda) - R_A(\mu) = (\lambda - \mu)R_A(\lambda)R_A(\mu)$$

(ii) *Pour λ, $\mu \in \rho(A)$, on a $R_A(\lambda)R_A(\mu) = R_A(\mu)R_A(\lambda)$.*

Démonstration. (i) L'équation de la résolvante se vérifie facilement ; jointe à la continuité de R_A, elle entraîne l'holomorphie de celle-ci, car

$$\lim_{\mu \to \lambda} \frac{R_A(\lambda) - R_A(\mu)}{\lambda - \mu} = R_A^2(\lambda)$$

De plus, par symétrie on a $R_A(\mu) - R_A(\lambda) = (\mu - \lambda)R_A(\mu)R_A(\lambda)$. Au vu de l'équation de la résolvante, on en déduit que $R_A(\lambda)$ et $R_A(\mu)$ commutent, ce qui démontre (ii). □

Théorème 3.3.6. *Le spectre de tout élément A de $\mathcal{L}(E)$ est un compact non vide de $\mathbb{C}$, inclus dans le disque fermé de centre 0 et de rayon $\|A\|$.*

Démonstration. L'application $\lambda \mapsto A - \lambda I$ est continue et $\rho(A)$ est l'image réciproque, par cette application, de l'ensemble des éléments de $\mathcal{L}(E)$ inversibles. Celui-ci étant un ouvert (lemme 3.3), on en déduit que $\rho(A)$ est un ouvert dans $\mathbb{C}$. D'autre part, tout $\lambda \in \mathbb{C}$, $|\lambda| > \|A\|$, est dans $\rho(A)$ car

$$A - \lambda I = -\lambda(I - \lambda^{-1}A)$$

et $\|\lambda^{-1}A\| < 1$. Il en résulte que $\sigma(A)$ est un fermé inclus dans le disque fermé $\{ \lambda \in \mathbb{C},\ |\lambda| \leq \|A\| \}$.

Il reste à vérifier que $\sigma(A)$ est non vide. Supposons le contraire ; pour x et y fixés,la fonction $\lambda \mapsto \langle R_A(\lambda)x, y\rangle$ est alors holomorphe sur $\rho(A) = \mathbb{C}$, et tend vers 0 quand $|\lambda| \to +\infty$; elle est donc bornée dans tout le plan complexe et le théorème de Liouville implique qu'elle est identiquement nulle

$$\langle R_A(\lambda)x, y\rangle = 0, \quad \forall\, x,\ y \in E$$

ceci nécessite que $R_A(\lambda) = 0.\ \forall\, \lambda \in \mathbb{C}$, ce qui est absurde. □

Théorème 3.3.7. *Soit A un élément de $\mathcal{L}(E)$ et soit p un polynôme de degré n, à coefficients complexes, $p(x) = \sum_0^n \alpha_k x^k$. Alors*

(i) *le spectre de $p(A)$ est $\sigma(p(A)) = p(\sigma(A)) = \{ p(\lambda) \mid \lambda \in \sigma(A) \}$*

(ii) *Si A est inversible, $\sigma(A^{-1}) = [\sigma(A)]^{-1} = \{ \lambda^{-1} \mid \lambda \in \sigma(A) \}$.*

Démonstration. Pour tout λ_0, $p(\lambda) - p(\lambda_0) = (\lambda - \lambda_0)q(\lambda)$, où q est aussi un polynôme. On en déduit que

$$p(A) - p(\lambda_0)I = (A - \lambda_0 I)q(A)$$

Si λ_0 appartient à $\sigma(A)$, alors $p(\lambda_0)$ appartient à $\sigma(p(A) - p(\lambda_0)I)$, sinon $[p(A) - p(\lambda_0)I]^{-1}q(A)$ serait l'inverse de $A - \lambda_0 I$. Ainsi, on a montré l'inclusion

$$p(\sigma(A)) \subset \sigma(p(A))$$

Inversement, soit λ_0 dans $\sigma(p(A))$ et désignons par λ_j, $1 \leq j \leq n$, les racines de $p(\lambda) - \lambda_0$, chacune étant répétée autant de fois que sa multiplicité. Il existe une constante α telle que

$$p(\lambda) - \lambda_0 = \alpha(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$$

et

$$p(A) - \lambda_0 I = \alpha(A - \lambda_1 I)\cdots(A - \lambda_n I)$$

Puisque le premier membre de cette dernière égalité est un opérateur non inversible, il existe j, $1 \leq j \leq n$, tel que $A - \lambda_j I$ est non inversible, c'est-à-dire $\lambda_j \in \sigma(A)$. En prenant $\lambda = \lambda_j$ dans la première égalité, on trouve $p(\lambda_j) = \lambda_0$, ce qui veut dire que $\lambda_0 \in p(\sigma(A))$ et par suite le spectre de $p(A)$ est inclus dans $p(\sigma(A))$. □

Par exemple, le spectre d'un opérateur nilpotent est réduit à 0 (un élément A de $\mathcal{L}(E)$ est nilpotent s'il existe n tel que $A^n = 0$).

Si $\lambda \in \mathbb{C}$ est tel que $A - \lambda I$ n'est pas injectif, alors λ appartient au spectre de A ; on dit que λ est une valeur propre de A. tout vecteur $x \neq 0$ appartenant au noyau de $A - \lambda I$ est appelé un vecteur propre de A correspondant à la valeur propre λ, ces vecteurs propres et 0 forment un sous-espace vectoriel fermé de E appelé sous-espace propre de A correspondant à la valeur propre λ, sa dimension (finie ou infinie) est la multiplicité de λ.

Définition 3.3.8. *Le spectre ponctuel de A, noté $\sigma_p(A)$, est l'ensemble des valeurs propres de A.*

Notons bien que, lorsque E est de dimension finie n, l'algèbre linéaire élémentaire nous a appris que toute valeur spectrale de A est une valeur propre de A, c'est-à-dire $\sigma(A) = \sigma_p(A)$, et le spectre de A est un ensemble fini ayant au plus n éléments qui sont les racines du polynôme caractéristique de A. Mais si E est de dimension infinie la situation se complique et on a en général seulement l'inclusion $\sigma_p(A) \subset \sigma(A)$. Naturellement, nous nous intéresserons aux cas "modérés" où la situation est "proche" de celle en dimension finie, à savoir les opérateurs dont le spectre est constitué d'un nombre (au plus) dénombrable de valeurs propres : C'est la classe des opérateurs *compacts* qui feront l'objet du chapitre IV. Voyons à présent, quelques exemples.

Exemple 3.3.9. - Déterminons le spectre de l'opérateur de translation défini sur $\ell^2(\mathbb{N})$ par $Ae_1 = 0$ et $Ae_{n+1} = e_n$. On peut vérifier facilement que

$\|A\| = 1$, si bien que pour tout complexe λ de module strictement supérieur à 1, l'opérateur

$$\lambda I - A = \lambda(I - \frac{A}{\lambda})$$

est inversible. Le spectre de A est donc inclus dans le disque unité fermé $\overline{D(0,1)}$. Nous allons démontrer qu'il lui est égal. On vérifie d'abord que pour tout λ, $|\lambda| < 1$, l'élément $x_\lambda = (1, \lambda, \lambda^2, \ldots)$ est dans $\ell^2(\mathbb{N})$ et satisfait la relation $Ax_\lambda - \lambda x_\lambda = 0$, donc $\lambda \in \sigma_p(A)$ et x_λ est un vecteur propre associé. On peut voir facilement que tout autre vecteur propre asssocié à λ est proportionnel à x_λ. Ainsi, on a montré l'inclusion de $D(0,1)$ dans $\sigma_p(A)$ et, compte tenu de ce qui précède, $\sigma(A) = \overline{D(0,1)}$. Par ailleurs, si $|\lambda| = 1$ la solution x_λ de $Ax = \lambda x$ n'appartient pas à $\ell^2(\mathbb{N})$, un tel λ n'appartient pas donc au spectre ponctuel ; comme il est dans $\sigma(A)$, forcément l'opérateur $A - \lambda I$ n'est pas surjectif. On peut d'ailleurs vérifier que l'unique solution $x = (x_1, x_2, \ldots)$ de l'équation $Ax - \lambda x = e_1$ est de la forme $x = x_1 e_1 + (1 + \lambda x_1)e_2 + \lambda(1 + \lambda x_1)e_3 + \cdots + \lambda^{n-2}(1 + \lambda x_1)e_n + \cdots$, et cet élément n'est pas dans $\ell^2(\mathbb{N})$.

Exemple 3.3.10. - Sur l'espace de Hilbert $L^2([0,1], dx)$, on considère l'opéra-teur A de multiplication : $Af(x) = xf(x)$. On a $\|A\| = 1$ et donc

$$\sigma(A) \subset \{\, \lambda \in \mathbb{C} \mid |\lambda| \leq 1 \,\}$$

Si λ est en dehors de l'intervalle $[0,1]$, la fonction $x \mapsto (x - \lambda)^{-1}$ est bornée sur l'intervalle $[0,1]$ et l'opérateur de multiplication par cette fonction est borné, c'est l'inverse de $A - \lambda I$. Donc $\sigma(A) \subset [0,1]$. D'autre part, la relation $Af - \lambda f = 0$ implique que f est nulle sauf éventuellement au point λ, l'opérateur A n'admet donc pas de valeur propre ($\sigma_p(A) = \emptyset$). Enfin, pour tout $\lambda \in [0,1]$, l'opérateur $A - \lambda I$ n'est pas surjectif, puisque l'unique solution de l'équation $Af - \lambda f = 1$ est la fonction $(x - \lambda)^{-1}$ qui n'appartient pas à $L^2([0,1], dx)$. En conclusion

$$\sigma_p(A) = \emptyset \quad \text{et} \quad \sigma(A) = [0,1].$$

Exemple 3.3.11. - Soit $\ell^2(\mathbb{N})$ l'espace de Hilbert des suites $\mathbf{x} = (x_n)$ de nombres complexes vérifiant

$$\sum_{n=0}^{\infty} |x_n|^2 < \infty$$

Soit D un compact de $\mathbb{C}$ et soit $\{\lambda_k, k \in \mathbb{N}\}$ une suite dense dans D. On désigne par A l'opérateur de multiplication par la suite (λ_k) :

$$\mathbf{x} = (x_n) \in \ell^2(\mathbb{N}) \mapsto A\mathbf{x} = (\lambda_n x_n)$$

On remarque d'abord que, la suite (λ_n) étant bornée, l'opérateur A est un endomorphisme continu de $\ell^2(\mathbb{N})$. D'autre part, pour tout n, λ_n appartient à $\sigma_p(A)$ puisque $Ae_n = \lambda_n e_n$. Comme le spectre de A est fermé, on en déduit que $D \subset \sigma(A)$. Soit $\lambda \in \mathbb{C} \setminus D$, il existe $\alpha > 0$ tel que $|\lambda - \lambda_k| > \alpha$, $\forall k$. La suite $((\lambda - \lambda_n)^{-1})$ est bornée et l'opérateur de multiplication par cette suite est donc borné, c'est l'inverse de $A - \lambda I$. Ainsi, un tel λ appartient à $\rho(A)$ et donc le spectre de A est inclus dans D. Finalement on a les égalités

$$\sigma(A) = D \quad \text{et} \quad \sigma_p(A) = \{ \lambda_n, n \in \mathbb{N} \}.$$

Cet exemple fournit, en quelque sorte, une *réciproque* du théorème 3.6, puisqu'il montre que tout compact de $\mathbb{C}$ est le spectre d'un opérateur linéaire continu sur un espace de Hilbert.

Exemple 3.3.12. - (spectre de l'opérateur de Volterra) : Soit $[a,b]$ un intervalle de $\mathbb{R}$, soit k une fonction continue sur $[a,b] \times [a,b]$ à valeurs réelles. Sur l'espace de Hilbert $E = L^2([a,b], dx)$, on considère l'opérateur de Volterra V défini par

$$Vf(x) = \int_a^x k(x,y) f(y)\, dy, \quad \text{pour } f \in L^2([a,b], dx)$$

On vérifie immédiatement que V^2 est un opérateur de Volterra et

$$V^2 f(x) = \int_a^x k_2(x,y) f(y)\, dy, \quad \text{avec} \quad k_2(x,y) = \int_y^x k(x,t) k(t,y)\, dt$$

Plus généralement, on vérifie par induction que pour tout entier $n \geq 2$, V^n est un opérateur de Volterra dont le noyau k_n est donné par

$$k_n(x,y) = \int_y^x k(x,t) k_{n-1}(t,y)\, dt, \quad y \leq x$$

Pour ces calculs et ceux qui suivent on peut voir l'exercice 7. Soit M le maximum du noyau k sur $[a,b] \times [a,b]$. On vérifie par récurrence sur n que

$$|k_n(x,y)| \leq M^n \frac{(x-y)^{n-1}}{(n-1)!} \leq M^n \frac{(b-a)^{n-1}}{(n-1)!}$$

De cette dernière inégalité il résulte que la série $\sum_{n=1}^{\infty} k_n(x,y)$ converge uniformément sur $[a,b] \times [a,b]$ vers une fonction continue sur $[a,b] \times [a,b]$. Ensuite, en utilisant l'inégalité de Cauchy-Schwarz, il vient

$$\begin{aligned} |V^n f(x)|^2 &\leq \Big(\int_a^x |k_n(x,t)|^2\, dt \Big) \|f\|_2^2 \\ &\leq \left(\frac{M^n}{(n-1)!} \right)^2 \frac{(x-a)^{2n-1}}{2n-1} \|f\|_2^2 \end{aligned}$$

Par intégration sur $[a, b]$, on en déduit que

$$\|V^n\| \leq M^n \frac{(b-a)^n}{(n-1)!\sqrt{2n}}$$

La série $\sum_{n=0}^{\infty} V^n$ est donc convergente dans $\mathcal{L}(E)$, ce qui implique que l'opérateur $I - V$ est inversible dans $\mathcal{L}(E)$. Le calcul précédent vaut pour l'opérateur $\lambda I - V = \lambda(I - \lambda^{-1}V)$, quel que soit $\lambda \neq 0$, il suffit simplement de remplacer M par $\lambda^{-1}M$. Ainsi, pour tout $\lambda \neq 0$, l'opérateur $\lambda I - V$ est inversible et donc $\sigma(V) = \{0\}$.
Notons que les calculs précédents montrent que l'inverse de $I - V$ est encore un opérateur de Volterra de noyau la fonction continue qui représente la somme de la série $\sum_{n=0}^{\infty} k_n(.,.)$.

EXERCICES

1. Sur l'espace de Hilbert $L^2([0,1], dx)$, on considère l'opérateur de multiplication défini par

$$Af(x) = m(x)f \quad \text{où} \quad m(x) = \begin{cases} 2x, & \text{si } 0 \leq x \leq \frac{1}{2}\,; \\ 1, & \text{si } frac12 \leq x \leq 1. \end{cases}$$

Montrer que $\sigma(A) = [0,1]$ et que $\sigma_p(A) = \{1\}$.
Quelle est la multiplicité de la valeur propre 1 ?

Solution : Soit λ dans $]0,1[$. La seule solution possible de l'équation $(Af - \lambda f) = 1$ est définie presque partout (en fait sauf en $x = \lambda/2$) par

$$f(x) = \begin{cases} \frac{1}{2x-\lambda} & \text{si } 0 \leq x \leq \frac{1}{2}\,; \\ \frac{1}{1-\lambda} & \text{si } \frac{1}{2} \leq x \leq 1. \end{cases}$$

Or, cette fonction n'appartient pas à $L^2([0,1], dx)$. Ainsi, pour un tel λ, l'opérateur $A - \lambda I$ n'est pas surjectif. Il en résulte que le spectre de A contient $]0,1[$, comme il est fermé, il est donc égale à $[0,1]$.
Soit λ dans $]0,1[$. L'équation $Af - \lambda f = 0$ implique que

$$(m(x) - \lambda)f(x) = 0, \quad \text{pour presque tout } x \in [0,1]$$

Compte tenu de l'expression de m, cela implique, nécessairement, que f est presque partout nulle. Pour $\lambda = 1$, la fonction $x \mapsto m(x) - 1$ est nulle sur $[1/2, 1]$. Il en résulte que, toute fonction de $L^2([0,1])$ à support dans $[1/2, 1]$ est fonction propre de A relativement à la valeur propre 1. Celle-ci est donc de multiplicité infinie. □

2. Soit E un espace de Hilbert séparable E et soit $(e_n)_{n \geq 1}$ une base hilbertienne de E. On considère l'opérateur A défini par $Ae_n = e_{n+1}$,

pour $n \geq 1$. Autrement dit, si $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$ est un élément arbitraire de E,

$$Ax = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_{n+1}$$

Déterminer $\sigma(A)$ et $\sigma_p(A)$.

Solution :

a) La définition de A montre que, pour tout $x \in E$, $\|Ax\| = \|x\|$. Il en résulte que $\|A\| = 1$, que A est injectif et que son spectre est inclus dans le disque fermée, $\overline{D(0,1)}$, de centre l'origine et de rayon 1.

b) Soit λ un nombre complexe non nul, avec $|\lambda| \leq 1$. L'équation $Ax - \lambda x = 0$ s'écrit

$$\sum_{n=1}^{\infty} \langle x, e_n \rangle e_{n+1} - \lambda \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n = 0$$

On en déduit que, pour tout entier $n \geq 1$, on a $\langle x, e_n \rangle = \lambda \langle x, e_{n+1} \rangle$. Par récurrence, il vient

$$\forall n \geq 0, \quad \langle x, e_{n+1} \rangle = \frac{1}{\lambda^n} \langle x, e_1 \rangle$$

Comme la série de terme général $1/\lambda^n$ n'appartient pas à $\ell^2(\mathbb{N})$, la solution x, ainsi trouvée, n'appartient pas à E. Il en résulte que λ n'est pas valeur propre de A. Le spectre ponctuel de A est donc réduit à l'ensemble vide.

c) Soit λ un nombre complexe non nul, avec $|\lambda| \leq 1$. L'équation $Ax - \lambda x = e_1$ s'écrit

$$\sum_{n=1}^{\infty} \langle x, e_n \rangle e_{n+1} - \lambda \sum_{n=1}^{\infty} \langle x, e_n \rangle = e_1$$

On en déduit que $\langle x, e_1 \rangle = -1/\lambda$ et par récurrence que

$$\forall n \geq 2, \quad \langle x, e_n \rangle = -\frac{1}{\lambda^n}$$

La série de terme général $-(1/\lambda^n)$ n'étant pas dans $\ell^2(\mathbb{N})$, on en déduit que l'équation $Ax - \lambda x = e_1$ n'admet pas de solution dans E, c'est-à-dire que l'opérateur $A - \lambda I$ n'est pas surjectif. Il en résulte que λ appartient au spectre de A.

En conclusion, on a $\sigma_p(A) = \emptyset$ et $\sigma(A) = \overline{D(0,1)}$. □

3. Soit a un réel non nul et soit τ_a l'opérateur de translation défini sur $L^2(\mathbb{R}, dx)$ par $\tau_a f(x) = f(x - a)$.

Montrer que τ_a est inversible et que $\|\tau_a\| = \|\tau_a^{-1}\| = 1$.
Montrer que $\sigma(\tau_a) \subset \{ \lambda \in \mathbb{C} \mid |\lambda| = 1 \}$.

Solurion : Pour toute f dans $L^2(\mathbb{R}, dx)$, on a $\tau_{-a}\tau_a f = \tau_a \tau_{-a} f = f$. Cela montre que τ_a est inversible et que $(\tau_a)^{-1} = \tau_{-a}$. D'autre part, la mesure de Lebesgue sur $\mathbb{R}$ étant invariante par translation, on en déduit que

$$\|\tau_a f\| = \|(\tau_a)^{-1} f\| = \|f\|$$

Les opérateurs τ_a et τ_{-a} sont donc injectifs, de norme 1 et leurs spectres sont inclus dans $\overline{D(0,1)}$. On déduit que $\sigma(\tau_a) \subset \overline{D(0,1)}$ et par suite, tout $\lambda \in \mathbb{C}$ de module $|\lambda| > 1$, est dans l'ensemble résolvant de $\tau_{\pm a}$. Soit λ un nombre complexe non nul, avec $|\lambda| < 1$. On a

$$\tau_a - \lambda I = -\lambda \tau_a (\tau_{-a} - \lambda^{-1} I)$$

Comme $|\lambda^{-1}| > 1$, le second membre de l'égalité ci-dessus est un opérateur inversible et il en est de même du premier. On en déduit qu'un tel λ est dans l'ensemble résolvant de $\tau_{\pm a}$. Finalement, le spectre de l'opérateur $\tau_{\pm a}$ est inclus dans le cercle unité. □

4. Soit $\phi \in L^1(\mathbb{R}, dx)$ et soit A l'opérateur de convolution par ϕ, défini sur l'espace $L^2(\mathbb{R}, dx)$ par

$$Af = \phi * f = \int_{\mathbb{R}} \phi(x-t) f(t)\, dt, \quad f \in L^2(\mathbb{R}, dx)$$

Montrer que le spectre de A est l'adhérence de l'ensemble des valeurs de la transformée de Fourier de ϕ.

Solution : Pour tout élément g de $L^2(\mathbb{R}, dx)$, on a l'équivalence

$$f \in L^2(\mathbb{R}, dx), \quad Af - \lambda f = g \Longleftrightarrow \hat{f} \in L^2(\mathbb{R}, dx), \quad (\hat{\phi} - \lambda)\hat{f} = \hat{g}$$

Si λ appartient à l'adhérence des valeurs de $\hat{\phi}$, la seule solution possible de l'équation $(\hat{\phi} - \lambda)\hat{f} = \hat{g}$, donnée par $\hat{f} = \hat{g}/(\hat{\phi} - \lambda)$ n'appartient pas à l'espace $L^2(\mathbb{R}, dx)$. L'opérateur $A - \lambda I$ n'est donc pas surjectif. Un tel λ est dans le spectre de A. En revanche, si λ est dans le complémentaire de l'adhérence des valeurs de $\hat{\phi}$, la fonction

$$\psi = \frac{1}{\hat{\phi} - \lambda}$$

est bien définie sur $\mathbb{R}$. De plus, la fonction ϕ étant dans $L^1(\mathbb{R}, dx)$, sa transformée de Fourier $\hat{\phi}$ est une fonction continue qui tend vers 0 à l'infini. On en déduit que la fonction ψ est continue et bornée sur $\mathbb{R}$. Il en résulte que la fonction $\psi \hat{g}$ est dans $L^2(\mathbb{R}, dx)$. Sa tranformée de Fourier inverse est alors l'unique solution de l'équation $Af - \lambda f = g$. Un tel λ est donc dans l'ensemble résolvant de l'opérateur A. □

5. Soit $\phi \in L^1(\mathbb{T}, dt)$ et soit A l'opérateur de convolution par ϕ, défini dans l'espace $L^2(\mathbb{T}, dt)$ par

$$Af = \frac{1}{2\pi} \int_0^{2\pi} \phi(x-t) f(t)\, dt$$

Montrer que le spectre de l'opérateur A est l'ensemble constitué des coefficients de Fourier de ϕ et de 0.

Solution : Pour $n \in \mathbb{Z}$ et $f \in L^2(\mathbb{T}, dt)$, on désigne par $c_n(f)$ les coefficients de Fourier de f. Soit λ non nul et vérifiant pour tout n : $\lambda \neq c_n(\phi)$. Pour g dans $L^2(\mathbb{T}, dt)$, l'équation $Af - \lambda f = g$ est équivalente à la relation $c_n(f) = c_n(g)/(c_n(\phi) - \lambda)$, pour tout entier $n \in \mathbb{Z}$. Comme $c_n(\phi)$ tend vers 0, quand n tend vers l'infini, il existe une constante M telle que pour tout n on ait $|c_n(\phi) - \lambda|^{-1} \leq M$. La suite $c_n(g)/(c_n(\phi) - \lambda)$ est donc dans $\ell^2(\mathbb{Z})$ et définit bien une fonction f de $L^2(\mathbb{T}, dt)$. Ainsi, tout λ non nul et distinct de $c_n(\phi)$ pour tout n est dans l'ensemble résolvant. D'autre part, f est fonction propre de A correspondant à une valeur propre λ (c'est-à-dire est solution de $Af - \lambda f = 0$) si, et seulement si, il existe $n \in \mathbb{Z}$ tel que $\lambda = c_n(\phi)$ et alors $f(x) = e^{inx}$. Cela veut dire que les valeurs propres de l'opérateur A sont les coefficients de Fourier de ϕ, comme le spectre de A est fermé et que $c_n(\phi)$ tend vers 0 quand n tend vers l'infini, il en résulte que $\sigma(A) = \{ c_n(\phi), n \in \mathbb{Z} \} \cup \{0\}$. □

6. On rappelle que le noyau de Poisson du disque unité est défini pour tout r, $0 \leq r < 1$, par

$$p_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}$$

et que

$$p_r(\theta) = \sum_{n \in \mathbb{Z}} r^{|n|} e^{in\theta}$$

Soit $\mathcal{P}_r$ l'opérateur de convolution associé à p_r :

$$\mathcal{P}_r f(\theta) = \frac{1}{2\pi} \int_0^{2\pi} f(\theta') p_r(\theta - \theta')\, d\theta', \quad \text{pour} \quad f \in L^2[0, 2\pi]$$

Montrer que les fonctions $e_n(\theta) = e^{in\theta}$, $n \in \mathbb{Z}$, sont des fonctions propres de $\mathcal{P}_r$ associées aux valeurs propres $\lambda_n = r^{|n|}$, $n \in \mathbb{Z}$.
Montrer que si $\lambda \neq \lambda_n$, $\forall n \in \mathbb{Z}$ et $\lambda \neq 0$, l'opérateur $\mathcal{P}_r - \lambda I$ est inversible.
En déduire que $\sigma(\mathcal{P}_r) = \{ 0, r^n, n \in \mathbb{N} \}$.

Solution : Le noyau de Poisson du disque unité est défni, pour tout r, $0 \leq r < 1$, par

$$p_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}$$

C'est une fonction continue, paire et 2π-périodique.
On vérifie rapidement que

$$\frac{1-r^2}{1-2r\cos\theta+r^2}=\frac{1-r^2}{1-r(e^{i\theta}+e^{-i\theta})+r^2}=\frac{1}{1-re^{i\theta}}+\frac{re^{-i\theta}}{1-re^{-i\theta}}$$

Comme $0<r<1$, le développement en série de chacun des deux derniers termes donne

$$p_r(\theta)=\sum_{n\in Z} r^{|n|}e^{in\theta}$$

Ainsi, pour tout $n\in\mathbb{Z}$, le coefficient de Fourier d'indice n de p_r est $r^{|n|}$. La fonction p_r étant paire, on a aussi

$$p_r(\theta-\theta'))=p_r(\theta'-\theta)=\sum_{n\in Z} r^{|n|}e^{-in\theta}e^{in\theta'}$$

Cela montre que la suite des coefficients de Fourier de la fonction $\theta'\mapsto p_r(\theta-\theta')$ est donc donnée par $n\mapsto r^{|n|}e^{-in\theta}$.
L'opérateur de convolution par la fonction p_r est défini, pour tout $f\in L^2[0,2\pi]$, par

$$\mathcal{P}_rf(\theta)=\frac{1}{2\pi}\int_0^{2\pi}f(\theta')p_r(\theta-\theta')d\theta'$$

Le second membre s'interprète comme le produit scalaire de f par la fonction $\theta'\mapsto p_r(\theta-\theta')$. La formule de Parseval permet alors d'en déduire que

$$\mathcal{P}_rf=\sum_{n\in\mathbb{Z}}\langle f,e_n\rangle r^{|n|}e_n$$

où l'on posé $e_n(\theta)=e^{in\theta}$. Pour que λ soit une valeur propre de $\mathcal{P}_r$, il faut et il suffit qu'il existe une fonction propre f (non identiquement nulle) telle que $\mathcal{P}_rf-\lambda f=0$, c'est-à-dire telle que

$$\sum_{n\in\mathbb{Z}}\langle f,e_n\rangle r^{|n|}e_n-\lambda\sum_{n\in\mathbb{Z}}\langle f,e_n\rangle e_n=0$$

ce qui s'écrit encore

$$\sum_{n\in\mathbb{Z}}\langle f,e_n\rangle(r^{|n|}-\lambda)e_n=0 \qquad (*)$$

Si, pour tout n dans $\mathbb{Z}$, λ est différent de $r^{|n|}$, l'équation $(*)$ implique que $\langle f,e_n\rangle=0$, $\forall n\in\mathbb{Z}$. Comme (e_n) est une base hilbertienne de E, on en déduit que f est nulle, ce qui est absurde. Un tel λ n'est donc

pas valeur propre de l'opérateur $\mathcal{P}_r$. En revanche, s'il existe un entier $k \in \mathbb{Z}$ tel que $\lambda = r^{|k|}$, alors l'équation (*) est satisfaite pour $f = e_k$ et pour $f = e_{-k}$. Cela montre que, pour tout $n \in \mathbb{Z}$, la fonction e_n est fonction propre de $\mathcal{P}_r$ et que la valeur propre associée est $\lambda_n = r^{|n|}$.
Soit $\lambda \neq 0$ et $\lambda \neq r^{|n|}, \forall n \in \mathbb{Z}$. L'opérateur $\mathcal{P}_r - \lambda I$ est injectif. Montrons qu'il est surjectif. Soit $g \in L^2[0, 2\pi]$ et cherchons s'il existe $f \in L^2[0, 2\pi]$ telle que

$$\mathcal{P}_r f - \lambda f = g$$

Compte tenu des calculs en 2, cette équation s'écrit

$$\sum_{n \in \mathbb{Z}} \langle f, e_n \rangle (r^{|n|} - \lambda) e_n = \sum_{n \in \mathbb{Z}} \langle g, e_n \rangle e_n$$

et nous en déduisons les relations

$$\forall n \in \mathbb{Z} \quad \langle f, e_n \rangle = \frac{1}{(r^{|n|} - \lambda)} \langle g, e_n \rangle$$

On vérifie que la suite $n \mapsto \langle f, e_n \rangle$ ainsi trouvée, est dans $\ell^2(\mathbb{Z})$. Il en résulte que la fonction

$$f = \sum_{n \in \mathbb{Z}} \frac{1}{(r^{|n|} - \lambda)} \langle g, e_n \rangle e_n$$

est bien dans $L^2[0, 2\pi]$. Ce qui prouve que l'opérateur $\mathcal{P}_r - \lambda I$ est surjectif et donc inversible.
Puisque le spectre de $\mathcal{P}_r$ est fermé, ce qui précède permet de conclure que : $\sigma_p(\mathcal{P}_r) = \{r^n, n \in \mathbb{N}\}$ et $\sigma(\mathcal{P}_r) = \{0, r^n, n \in \mathbb{N}\}$. □

7. Soit V l'opérateur de Volterra défini sur $L^2([0,1], dx)$ par

$$Vf(x) = \int_0^x f(t)\, dt$$

Montrer que pour tout entier $n \geq 1$, on a

$$V^n f(x) = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)\, dt \tag{3.1}$$

En déduire que la résolvante $R_V(\lambda) = (A - \lambda I)^{-1}$, définie pour $\lambda \neq 0$, est donnée par

$$R_V(\lambda) f(x) = -\lambda^{-1} f(x) - \lambda^{-2} \int_0^x e^{\lambda^{-1}(x-t)} f(t)\, dt \tag{3.2}$$

Solution : L'expression de $V^n f(x)$ se démontre par récurrence. Elle est vraie pour $n = 1$ et si on suppose qu'elle l'est pour un entier n, il

vient

$$V^{n+1}f(x) = \int_0^x V^n f(y)\,dy = \int_0^x \left(\int_0^y \frac{(y-t)^{n-1}}{(n-1)!} f(t)\,dt\right) dy$$
$$= \int_0^x \left(\int_t^x \frac{(y-t)^{n-1}}{(n-1)!} dy\right) f(t)\,dt = \int_0^x \frac{(x-t)^n}{n!} f(t)\,dt$$

Ce qui prouve la relation (1) pour $n+1$ à la place de n. On sait que le spectre de V est réduit à l'élément 0. Pour $\lambda \neq 0$, $V - \lambda I$ est inversible, pour trouver l'expression de $R_V(\lambda) = (V - \lambda I)^{-1}$, on écrit

$$(V - \lambda I)^{-1} f(x) = -\lambda^{-1}(I - \lambda^{-1}V)^{-1} f(x)$$
$$= -\lambda^{-1}\sum_{n=0}^{\infty} \lambda^{-n} V^n f(x), \quad (\text{avec} \quad V^0 = I)$$
$$= -\lambda^{-1} f(x) - \sum_{n=1}^{\infty} \lambda^{-n} \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)\,dt$$
$$= -\lambda^{-1} f(x) - \lambda^{-2} \sum_{m=0}^{\infty} \int_0^x \frac{\lambda^{-m}(x-t)^m}{m!} f(t)\,dt$$

En intervertissant les signes somme et intégrale, opération qu'on peut justifier, on arrive à l'expression désirée. □

8. Soit E un espace de Hilbert avec $E = E_1 \oplus E_2$, où E_1 et E_2 sont deux sous-espaces fermés dans E. Soit $A \in \mathcal{L}(E)$ qui laisse invariant chacun de ces deux sous-espaces, c'est-à-dire qui vérifie $AE_i \subset E_i$, $i = 1, 2$. Soit A_i la restriction de A à E_i. Montrer que le spectre de A est la réunion des spectres de A_1 et A_2.

   Solution : Il est clair que si $A - \lambda I$ est injectif, il en sera de même des opérateurs $A_i - \lambda I$, $i = 1, 2$. Inversement, si les deux opérateurs $A_i - \lambda I$, $i = 1, 2$, sont injectifs, il en sera de même de l'opérateur $A - \lambda I$. En effet, supposons qu'il exite $f = f_1 + f_2$ (avec $f_i \in E_i$) non identiquement nul tel que $Af - \lambda f = 0$. Il en résulte que $A_1 f_1 - \lambda f_1 = 0$ et $A_2 f_2 - \lambda f_2 = 0$. Mais, l'un au moins de f_1 et f_2 est non identiquement nul, et par suite l'un au moins des opérateurs $A_i - \lambda I$, $i = 1, 2$, est non injectif; ce qui contredit l'hypothèse. Ainsi, on a montrer que $A - \lambda I$ est injectif si, et seulement si, les deux opérateurs $A_i - \lambda I$, $i = 1, 2$, le sont. On montre de même que $A - \lambda I$ est surjectif si, et seulement si, les deux opérateurs $A_i - \lambda I$, $i = 1, 2$, le sont. □

3.4 Opérateur adjoint–Opérateur autoadjoint

Théorème 3.4.1. *Soient E un espace de Hilbert et $A \in \mathcal{L}(E)$. Il existe un unique opérateur continu de E dans E, noté A^* et appelé l'adjoint de A,*

tel que

$$\langle Ax, y\rangle = \langle x, A^*y\rangle, \quad \forall x,\ y \in E$$

En outre, on a $(A^*)^* = A$ *et* $\|A^*\| = \|A\| = \|A^*A\|^{\frac{1}{2}}$.

Démonstration. Pour tout y dans E, l'application qui à x associe $\langle Ax, y\rangle$ est une forme linéaire continue sur E, dont la norme n'excède pas $\|A\|\ \|y\|$; en vertu du théorème de représentation de Riesz, il existe un unique $z \in E$ tel que pour tout x dans E

$$\langle Ax, y\rangle = \langle x, z\rangle \quad \text{et} \quad \|z\| \leq \|A\|\ \|y\|$$

Posons $z = A^*y$, on montre aussitôt, en fixant x, que $y \to A^*y$ est une application linéaire et $\|A^*y\| \leq \|A\|\ \|y\|$, donc $\|A^*\| \leq \|A\|$. L'opérateur $(A^*)^*$ est caractérisé par la propriété :

$$\langle A^*x, y\rangle = \langle x, (A^*)^*y\rangle, \quad \text{pour} \ \ x,\ y \in E$$

Or,

$$\langle A^*x, y\rangle = \overline{\langle y, A^*x\rangle} = \overline{\langle Ay, x\rangle} = \langle x, Ay\rangle$$

On en déduit immédiatement que $(A^*)^* = A$ et par suite $\|A^*\| = \|A\|$. De plus, l'inégalité

$$\|Ax\|^2 \leq \|A^*A\|\ \|x\|^2, \quad \forall\ x \in E$$

montre que $\|A\|^2 \leq \|A^*A\| \leq \|A^*\|\ \|A\|$; comme $\|A^*\| = \|A\|$, on a l'égalité $\|A^*A\| = \|A\|^2$. □

Propriétés de l'adjoint *Soient A et B deux éléments de $\mathcal{L}(E)$. On a*

(a) $(\lambda A + \mu B)^* = \overline{\lambda} A^* + \overline{\mu} B^*$, *pour* $\lambda,\ \mu \in \mathbb{C}$
(b) $(AB)^* = B^*A^*$
(c) *Si A est inversible, A^* l'est aussi et on a* $(A^*)^{-1} = (A^{-1})^*$
(d) $e^{A^*} = (e^A)^*$
(e) $\sigma(A^*) = \overline{\sigma(A)} = \{\ \overline{\lambda} \mid \lambda \in \sigma(A)\ \}$

Démonstration. Les propriétés (a), (b) et (d) sont faciles à prouver, alors que (c) est une conséquence de (b), qui assure que $(AA^{-1})^* = (A^{-1})^*A^* = A^*(A^{-1})^* = I$, et du fait que $I^* = I$. Enfin (e) est une conséquence directe de (c). □

Exemple 3.4.2. - Soit (X, Ω, μ) un espace mesuré σ-fini, soit ϕ une fonction μ-mesurable essentiellement bornée et A_ϕ l'opérateur de multiplication par ϕ dans l'espace de Hilbert $L^2(X, \Omega, \mu)$. Alors l'adjoint A_ϕ^* de A_ϕ est l'opérateur de multiplication par $\overline{\phi}$, c'est-à-dire que $A_\phi^* = A_{\overline{\phi}}$.

• Si un opérateur sur $\mathbb{K}^n$ est représenté par sa matrice, alors son adjoint est représenté par la matrice transposée conjuguée.

• Soit E est un espace de Hilbert séparable et soit (e_n) une base hilbertienne de E. On a vu dans la section 1 que tout élément A de $\mathcal{L}(E)$ est représenté par la *matrice infinie* (a_{ij}), avec $a_{ij} = \langle Ae_j, e_i\rangle$. On peut voir de même que l'opérateur adjoint A^* est représenté par la matrice (a^*_{ij}) donnée par $a^*_{ij} = \langle A^*e_j, e_i\rangle$. Comme

$$\langle A^*e_j, e_i\rangle = \langle e_j, Ae_i\rangle = \overline{\langle Ae_i, e_j\rangle} = \overline{a_{ji}}$$

on en déduit, comme en dimension finie, que l'adjoint A^* est représenté par la matrice transposée conjuguée.

Exemple 3.4.3. - Désignons par (e_n) la base canonique de $\ell^2(\mathbb{N})$ et soit A l'opérateur de translation défini sur $\ell^2(\mathbb{N})$ par

$$rlAe_1 = 0, \quad Ae_{n+1} = e_n, \quad n \geq 1$$

Les relations

$$\langle A^*e_j, e_n\rangle = \langle e_j, Ae_n\rangle = \begin{cases} 1, & \text{si } j+1 = n\,; \\ 0, & \text{sinon.} \end{cases}$$

montrent que A^* est défini par $A^*e_j = e_{j+1}$, $j \geq 1$. Il en résulte que, si $x = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$ est un élément arbitraire de $\ell^2(\mathbb{N})$, ses images par A et par A^* sont données par

$$Ax = \sum_{n=2}^{\infty} \langle x, e_n\rangle e_{n-1}, \quad A^*x = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_{n+1}$$

Exemple 3.4.4. - Soient un intervalle fermé borné $[a,b]$, une fonction k continue sur $[a,b] \times [a,b]$ à valeurs complexes et K l'opérateur intégral de Fredholm de noyau k défini, pour $f \in L^2([a,b], dx)$, par

$$Kf(x) = \int_a^b k(x,y) f(y)\, dy$$

C'est un opérateur linéaire borné sur E. Pour f et g dans E, on a

$$\begin{aligned}\langle Kf, g\rangle &= \int_a^b \left(\int_a^b k(x,y) f(y)\, dy\right) \overline{g(x)}\, dx \\ &= \int_a^b f(y) \left(\int_a^b k(x,y) \overline{g(x)}\, dx\right) dy \\ &= \int_a^b f(y) \overline{\left(\int_a^b \overline{k(x,y)} g(x)\, dx\right)}\, dy = \langle f, K^*g\rangle\end{aligned}$$

On en déduit que K^* est l'opérateur intégral de Fredholm de noyau k^*, avec $k^*(x,y) = \overline{k(y,x)}$. Autrement dit

$$K^*f(x) = \int_a^b \overline{k(y,x)} f(y)\, dy$$

Le théorème suivant est souvent utilisé :

Théorème 3.4.5. *Soit $A \in \mathcal{L}(E)$. Pour que A soit injectif, il faut et il suffit que l'image de A^* soit dense dans E, ce qui se traduit par l'égalité*

$$(\Im m A^*)^{\perp} = \ker A$$

On remarque que l'égalité $(A^*)^* = A$ permet de remplacer, dans cet énoncé, A par A^*. Autrement dit, pour que l'image de A soit dense dans E, il faut et il suffit que A^* soit injectif, c'est-à-dire que $(\Im m A)^{\perp} = \ker A^*$.

Démonstration. Pour que y soit orthogonal à l'image de A, il faut et il suffit que $\langle Ax, y\rangle = \langle x, A^*y\rangle$ soit nul pour tout x dans E, ce qui implique que $A^*y = 0$. Ainsi, l'orthogonal de $A(E)$ est bien $\ker(A^*)$. La deuxième relation s'en déduit en remplaçant A par A^*. □

Définition 3.4.6. *Un opérateur $A \in \mathcal{L}(E)$ est dit auto-adjoint (ou parfois hermitien) s'il est égal à son adjoint, c'est-à-dire si, quels que soient x et y dans E, $\langle Ax, y\rangle = \langle x, Ay\rangle$.*

Exemple 3.4.7. -

(1) Un opérateur intégral de Fredholm est auto-adjoint si et seulement si son noyau k vérifie $\overline{k(y,x)} = k(x,y)$

(2) L'opérateur de multiplication par une fonction $\phi \in L^{\infty}(X, \Omega, \mu)$ est auto-adjoint dans $L^2(X, \Omega, \mu)$, si et seulement si ϕ est à valeurs réelles.

(3) Soit $A \in \mathcal{L}(E)$. Les opérateurs $U = \frac{1}{2}(A + A^*)$ et $V = \frac{1}{2i}(A - A^*)$ sont auto-adjoints et on a $A = U + iV$. L'opérateur U est appelé *la partie réelle* de A et l'opérateur V est sa *partie imaginaire*.

Le résultat suivant, outre qu'il fournit des exemples d'opérateurs auto-adjoints, est un outil très utile dans la discussion de l'existence de solutions d'équations aux dérivées partielles elliptiques. Il peut être considéré comme une variante du théorème de représentation de Riesz (théorème 3.4, chapitre I).

Théorème 3.4.8. (Lax-Milgram) *Soit E un espace de Hilbert et soit B une forme hermitienne sur $E \times E$. On suppose que :*

(1) *la forme B est continue, c'est-à-dire qu'il existe une constante M telle que, pour x et y dans E on ait $|B(x,y)| \le M\|x\|\,\|y\|$*

(2) *la forme B est coercitive, c'est-à-dire qu'il existe une constante $C > 0$ telle que, pour tout $x \in E$, $B(x,x) \ge C\|x\|^2$*

Alors, il existe un unique opérateur A dans $\mathcal{L}(E)$, inversible et auto-adjoint , tel que

$$B(x,y) = \langle Ax, y\rangle, \quad \|A\| \le M \quad \text{et} \quad \|A^{-1}\| \le C.$$

Démonstration. Pour y fixé dans E l'application qui à x associe $B(x,y)$ est une forme linéaire continue sur E dont la norme est majorée par $M\|y\|$ car

$$|B(x,y)| \le M\|y\|\,\|x\|, \quad \forall\, x \in E$$

D'après le théorème de représentation de Riesz, il existe un unique élément, qui dépend de y et que nous notons donc Ay, tel que

$$B(x,y) = \langle x, Ay\rangle$$

L'application A est linéaire car, pour x, y_1 et y_2 dans E et pour α dans $\mathbb{K}$, on a

$$\begin{aligned}\langle x, A(y_1+\alpha y_2)\rangle &= B(x, y_1+\alpha y_2) = B(x,y_1) + \overline{\alpha}B(x,y_2)\\ &= \langle x, Ay_1\rangle + \overline{\alpha}\langle x, Ay_2\rangle = \langle x, Ay_1 + \alpha Ay_2\rangle\end{aligned}$$

L'application A est bornée par M, puisque $|\langle x, Ay\rangle| \leq M\|x\|\,\|y\|$. Enfin A est auto-adjoint car, B étant hermitienne, la relation $\overline{B(x,y)} = B(y,x)$ se traduit par

$$\overline{\langle x, Ay\rangle} = \langle y, Ax\rangle$$

et le premier membre de cette égalité est précisément égal à $\langle Ay, x\rangle$. D'autre part, si $Ax = 0$ alors $B(x,x) = 0$, comme la forme B est coercitive, on en déduit que $x = 0$, ainsi $\ker A = \{0\}$ et le théorème 4.5 implique que

$$(\mathfrak{Im} A)^\perp = \{0\}$$

L'opérateur A est donc injectif à image dense. Montrons que son image est fermée, ce qui prouvera que A est surjectif et donc inversible. Soit (z_n) une suite de Cauchy dans l'image de A ; pour tout n il existe $x_n \in E$ tel que $z_n = Ax_n$. La coercitivité de la forme B implique de nouveau que

$$C\|x_n - x_m\| \leq \|Ax_n - Ax_m\|$$

Le second membre de cette inégalité tend vers 0 quand n et m tendent vers l'infini, donc (x_n) est une suite de Cauchy et si x est sa limite on aura

$$\lim_{n\to\infty} Ax_n = Ax$$

ce qui montre que la limite de la suite (z_n) appartient à l'image de A. Enfin, la coercitivité appliquée à $x = A^{-1}y$, $y \in E$, donne

$$C\|A^{-1}y\|^2 \leq \langle AA^{-1}y, A^{-1}y\rangle = \langle y, A^{-1}y\rangle \leq \|y\|\,\|A^{-1}y\|$$

d' où il résulte que $\|A^{-1}\| \leq C^{-1}$. □

Proposition 3.4.9. *Soit E un espace de Hilbert sur $\mathbb{C}$. Un élément A de $\mathcal{L}(E)$ est auto-adjoint si, et seulement si, pour tout x dans E, $\langle Ax, x\rangle$ est un nombre réel.*

Démonstration. Si A est auto-adjoint, alors $\langle Ax, x\rangle = \langle x, Ax\rangle$ et ce nombre est donc réel. Supposons maintenant que, pour tout élément x de E, $\langle Ax, x\rangle$ soit réel. Cela implique que

$$\langle Ax, x\rangle = \langle x, Ax\rangle, \quad \text{et par suite} \quad \langle (A - A^*)x, x\rangle = 0, \quad \forall\, x \in E$$

Ainsi, la forme sesquilinéaire qui à (x, y) associe $\langle (A - A^*)x, y\rangle$ s'annule en tout point de la diagonale de $E \times E$ et l'identité de polarisation (voir le chapitre I, 1.2) montre alors qu'elle s'annule en tout point (x, y) de $E \times E$, c'est-à-dire $\langle (A - A^*)x, y\rangle = 0$ quels que soient x et y dans E. Il s'ensuit que $A = A^*$. □

Corollaire 3.4.10. *Soient E un espace de Hilbert sur $\mathbb{C}$ et A dans $\mathcal{L}(E)$. Si $\langle Ax, x\rangle = 0$ pour tout $x \in E$, alors $A = 0$.*

Remarque 3.4.11. - La proposition 4.9 et le corollaire 4.10 tombent en défaut si E est supposé un espace de Hilbert sur $\mathbb{R}$, comme le montre l'exemple suivant

$$E = \mathbb{R}^2, \qquad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

On a $\langle Ax, x\rangle = 0$ pour tout $x \in \mathbb{R}^2$. Mais $A^* \neq A$. La raison est que l'identité de polarisation, dans le cas réel, n'est vraie que si la forme sesquilinéaire qui à (x, y) associe $\langle Ax, y\rangle$ est symétrique, ce qui veut dire précisément que A est auto-adjoint. On comprend donc les hypothèses du corollaire suivant

Corollaire 3.4.12. *Soient E un espace de Hilbert sur $\mathbb{R}$ et A un élément de $\mathcal{L}(E)$ auto-adjoint. Si $\langle Ax, x\rangle = 0$ pour tout $x \in E$, alors $A = 0$.*
Nous avons déj à à notre disposition plusieurs formules qui permettent de calculer la norme d'un opérateur $A \in \mathcal{L}(E)$ (voir le début de ce chapitre). Lorsque celui-ci est auto-adjoint, le calcul peut être considérablement simplifié gâce au théorème suivant

Théorème 3.4.13. *Si A est un opérateur auto-adjoint, alors*

$$\|A\| = \sup_{\|x\|=1} |\langle Ax, x\rangle|$$

De plus, si le maximum est atteint en x_0, alors celui-ci est vecteur propre de A associé à $\|A\|$ ou $-\|A\|$.

Démonstration. Désignons par $N(A)$ le maximum de $|\langle Ax, x\rangle|$ lorsque x varie sur la sphère unité. L'inégalité de Cauchy-Schwarz montre que $|\langle Ax, x\rangle|$ est inférieur à $\|A\|\ \|x\|^2$, on en déduit immédiatement l'inégalité $N(A) \leq \|A\|$. Pour montrer l'inégalité dans l'autre sens, on remarque d'abord que, par définition de $N(A)$, on a

$$|\langle Ax, x\rangle| \leq N(A)\|x\|^2, \quad \forall\, x \in E$$

d'autre part, en utilisant le fait que $A = A^*$, on montre facilement que $4\Re e\langle Ax, y\rangle = \langle A(x+y), (x+y)\rangle - \langle A(x-y), (x-y)\rangle$, pour x et y dans E. L'inégalité précédente et l'identité du parallélogramme donnent alors

$$ll4\Re e\langle Ax, y\rangle \leq N(A)\big(\|x+y\|^2 + \|x-y\|^2\big)$$
$$= 2N(A)\big(\|x\|^2 + \|y\|^2\big)$$

et par suite, quels que soient x et y de norme 1, $\Re e\langle Ax, y\rangle \leq N(A)$. Prenons $y = Ax/\|Ax\|$, lorsque $Ax \neq 0$, on en déduit que

$$\frac{\Re e\langle Ax, Ax\rangle}{\|Ax\|} = \|Ax\| \leq N(A)$$

En prenant le maximum sur l'ensemble des $x \in E$ de norme 1, on arrive à l'inégalité $\|A\| \leq N(A)$ et par suite à l'égalité cherchée.

Supposons maintenant que le maximum soit atteint en un point $x_0 \in E$ et soit z un élément de E de norme 1 et orthogonal à x_0. L'élément $x_t = (\cos t)x_0 + (\sin t)z$ est de norme 1 quel que soit t réel, et coïncide avec x_0 pour $t = 0$, il en résulte que la fonction $t \mapsto \langle Ax_t, x_t\rangle$ présente un extremum en $t = 0$ et par suite sa dérivée est nulle en ce point. Le calcul de cette dérivée donne

$$2\Re e(\langle Ax_0, z\rangle) = 0$$

En substituant iz à z, on voit que $\langle Ax_0, z\rangle = 0$. Ceci étant pour tout z de norme 1 et orthogonal à x_0, il en résulte que Ax_0 appartient à l'orthogonal de $\{x_0\}^\perp$, c'est-à-dire qu'il existe une constante λ telle que $Ax_0 = \lambda x_0$. On en déduit alors l'égalité

$$\lambda = \langle \lambda x_0, x_0\rangle = \langle Ax_0, x_0\rangle = \pm\|A\|$$

Ce qui termine la preuve. □

Il faut bien noter que le maximum n'est pas forcément atteint et donc les nombres $\pm\|A\|$ ne sont pas forcément des valeurs propres. Nous verrons dans le chapitre suivant que si l'opérateur A est compact alors le maximum est toujours atteint et donc $\|A\|$ ou $-\|A\|$ est valeur propre de A.

Proposition 3.4.14. *Soit $A \in \mathcal{L}(E)$ un opérateur auto-adjoint et soit F un sous-espace de E. Si F est invariant par A, alors $F^\perp$ est aussi invariant par A.*

Démonstration. En effet, pour tout $x \in F$ et $y \in F^\perp$, on a

$$\langle Ay, x\rangle = \langle y, Ax\rangle = 0$$

car Ax appartient à F, il en résulte que Ay appartient à $F^\perp$. □

Définition 3.4.15. *Un opérateur A de $\mathcal{L}(E)$ est dit positif si, pour tout $x \in E$, $\langle Ax, x\rangle$ est supérieur ou égal à zéro.*

Notons que tout opérateur positif, dans un espace de Hilbert sur $\mathbb{C}$, est auto-adjoint (proposition 4.10). D'autre part, pour tout opérateur $A \in \mathcal{L}(E)$, les opérateurs AA^* et A^*A sont auto-adjoints positifs car, pour tout x dans E, $\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle \geq 0$. En particulier si A est auto-adjoint, alors A^2 est un opérateur positif.

Théorème 3.4.16. *Le spectre d'un opérateur auto-adjoint est réel et, celui d'un opérateur auto-adjoint positif est positif. Les sous-espaces propres correspondants à deux valeurs propres distinctes d'un opérateur auto-adjoint, sont orthogonaux.*

Démonstration. Soit $A \in \mathcal{L}(E)$ un opérateur auto-adjoint et soit $\lambda = \mu + i\nu$, avec $\nu \neq 0$. Pour tout $x \in E$, on a

$$\|Ax - \lambda x\|^2 = \|Ax - \mu x\|^2 + \nu^2\|x\|^2$$

Il en résulte que, d'une part $A - \lambda I$ est un opérateur injectif, et que d'autre part son image est un sous-espace fermé de E. Il en est de même de l'opérateur $A - \overline{\lambda} I$. Utilisant le théorème 4.6 on en déduit que $\Im m(A - \lambda I)^{\perp} = \ker(A - \overline{\lambda} I) = \{0\}$ et par suite $\Im m(A - \lambda I) = E$. L'opérateur $A - \lambda I$ est donc bijectif de E sur E. Ainsi, tout $\lambda \in \mathbb{C} \backslash \mathbb{R}$ est dans l'ensemble résolvant de A. Si de plus l'opérateur A est positif, on vérifie que, pour tout $x \in E$ et tout $\lambda < 0$,

$$\|Ax - \lambda x\|^2 = \|Ax\|^2 - 2\lambda\langle Ax, x\rangle + \lambda^2\|x\|^2 \geq \lambda^2\|x\|^2$$

et on en déduit, comme précédemment, que $A - \lambda I$ est inversible.

Soient, maintenant, λ_1 et λ_2 deux valeurs propres distinctes de A et soient x_1 et x_2 deux vecteurs propres associés. Ces valeurs propres sont nécessairement réelles, d'après ce qui précède, et on a

$$(\lambda_1 - \lambda_2)\langle x_1, x_2\rangle = \langle Ax_1, x_2\rangle - \langle x_1, Ax_2\rangle = 0$$

Comme $\lambda_1 \neq \lambda_2$, on a nécessairement $\langle x_1, x_2\rangle = 0$. □

EXERCICES

1. Soit E un espace de Hilbert et A un élément de $\mathcal{L}(E)$. Montrer que si une suite (x_n) converge faiblement vers x dans E, alors la suite (Ax_n) converge faiblement vers Ax.
   En déduire que si (x_n) converge faiblement vers x et si (y_n) converge vers y, alors la suite $(\langle Ax_n, y_n\rangle)$ converge vers $\langle Ax, y\rangle$.

Solution : 1) Soit (x_n) une suite d'éléments de E qui converge faiblement vers x. Pour tout $y \in E$, on a

$$\lim_{n\to\infty} \langle Ax_n, y\rangle = \lim_{n\to\infty} \langle x_n, A^*y\rangle = \langle x, A^*y\rangle = \langle Ax, y\rangle$$

2) Supposons que (x_n) converge faiblement vers x et que (y_n) converge (fortement) vers y, on a

$$|\langle Ax, y\rangle - \langle Ax_n, y_n\rangle| \leq |\langle Ax, y\rangle - \langle Ax_n, y\rangle| + |\langle Ax_n, y\rangle - \langle Ax_n, y_n\rangle|$$

La première étape montre que le premier terme du second membre tend vers 0 quand n tend vers l'infini. D'autre part, puisque (x_n) converge faiblement elle est bornée, il existe donc M telle que $\|x_n\| \leq M$, pour tout entier n ; il en résulte que

$$|\langle Ax_n, y\rangle - \langle Ax_n, y_n\rangle| \leq M\|A\|\|y - y_n\|$$

Le second membre de cette inégalité tend vers 0 quand n tend vers l'infini. □

2. On considère l'opérateur de Volterra défini sur $L^2([0,1], dx)$ par

$$Vf(x) = \int_0^x f(t)\, dt$$

Déterminer l'adjoint de V.

Solution : Soient f et g deux éléments de $L^2([0,1], dx)$, qu'on peut supposer à valeurs réelles. On a

$$\langle Vf, g\rangle = \int_0^1 \left(\int_0^x f(t)\, dt\right) g(x)\, dx = \int_0^1 f(t) \left(\int_t^1 g(x)\, dx\right) dt$$

Cela montre que l'adjoint de l'opérateur V est donné par

$$V^*f(x) = \int_x^1 f(t)\, dt$$

3. (*Hellinger-Toeplitz.*) Soit E un espace de Hilbert sur $\mathbb{C}$. Soit A un endomorphisme de E hermitien, c'est-à-dire vérifiant

$$\langle Ax, y\rangle = \langle x, Ay\rangle, \quad \forall\, x, y \in E$$

Montrer que nécessairement A est continu sur E.

Solution : C'est une application directe du théorème du *graphe fermé.* Voici une autre manière de voir : puisque pour tout x dans E de norme 1, on a

$$|\langle Ax, y\rangle| = |\langle x, Ay\rangle| \leq \|Ay\|, \quad \forall\, y \in E$$

on en déduit que $\{ Ax;\ \|x\| \leq 1 \}$ est un ensemble faiblement borné. Le théorème de la borne uniforme assure alors que cet ensemble est borné, c'est à dire que

$$\sup\{ \|Ax\|;\ \|x\| \leq 1 \} = \|A\| < \infty.$$

4. Sur l'espace de Hilbert $L^2[0,1]$, on définit l'opérateur de multiplication A_ϕ, pour $\phi \in L^\infty[0,1]$, par

$$A_\phi f = \phi f$$

(i) Montrer que $\|A_\phi\| = \|\phi\|_\infty$.
(ii) Montrer que λ est une valeur propre de A_ϕ si, et seulement si, l'ensemble $\{x \mid \phi(x) = \lambda\}$ est de mesure non nulle.

5. Soit E un espace de Hilbert sur $\mathbb{C}$. Montrer que si $A \in \mathcal{L}(E)$ est un opérateur positif, c'est-à-dire vérifiant $\langle Ax, x\rangle \geq 0$, $\forall x \in E$, alors, quels que soient x et y dans E,

$$|\langle Ax, y\rangle|^2 \leq \langle Ax, x\rangle.\langle Ay, y\rangle$$

C'est une sorte de généralisation de l'inégalité de Cauchy-Schwarz.

Indication : Considérer $\langle A(x + \lambda y), x + \lambda(x + y)\rangle$, pour $\lambda \in \mathbb{R}$.

6. Soit E un espace de Hilbert. Pour deux éléments A et B dans $\mathcal{L}(E)$, on écrit $A \leq B$ si l'opérateur $B - A$ est positif. Soit (A_n) une suite croissante d'opérateurs auto-adjoints dans $\mathcal{L}(E)$, telle que

$$0 \leq A_1 \leq A_2 \leq \ldots \leq A_n \leq \ldots \leq I$$

(a) Soit $A_{ij} = A_j - A_i$, pour $i < j$. Montrer, en utilisant l'exercice précédent, que

$$\|A_{ij}x\|^4 \leq \|A_{ij}\|^3 \|x\|^2 \langle A_{ij}x, x\rangle$$

En déduire que

$$\|A_j x - A_i x\|^4 \leq \Big[\langle A_j x, x\rangle - \langle A_i x, x\rangle\Big] \|x\|^2$$

(b) En remarquant que $(\langle A_j x, x\rangle)$ est une suite croissante et bornée, montrer que $(A_j x)$ est une suite de Cauchy dans E. Soit Ax sa limite, montrer que l'opérateur A, ainsi défini, est linéaire borné et auto-adjoint.

7. Cet exercice renvoie au théorème de Lax-Milgram. Soit E un espace de Hilbert et soit B une forme hermitienne sur $E \times E$, coercitive.
(a) Montrer que pour toute forme linéaire L, continue sur E, il existe un unique élément $u \in E$ tel que

$$(*) \qquad L(v) = B(v, u), \quad \forall\, v \in E$$

(b) Soit ϕ l'application définie sur E par

$$\phi(v) = B(v, v) - 2\Re e(L(v))$$

Montrer que u est solution de $(*)$ si, et seulement si

$$\phi(u) \leq \phi(v), \quad \forall\, v \in E$$

8. Soit E un espace de Hilbert et A un élément de $\mathcal{L}(E)$ auto-adjoint.
(a) Montrer que si λ appartient à $\rho(A)$ alors il existe $c > 0$ telle que : $\|Ax - \lambda x\| \geq c\|x\|$, pour tout $x \in E$
(b) Soit $\mu \in \mathbb{R}$, et supposons qu'il existe une constante $c' > 0$ telle que : $\|Ax - \mu x\| \geq c'\|x\|$, pour tout $x \in E$. Montrer que μ n'est pas valeur propre de A et que l'image de l'opérateur $A - \mu I$ est dense dans E. Montrer que l'image de $A - \mu I$ est fermée et en déduire que μ appartient à $\rho(A)$.
(c) Déduire de ce qui précède qu'une condition nécessaire et suffisante pour qu'un nombre réel λ appartienne au spectre de A est qu'il existe une suite (x_n), $n \in \mathbb{N}$, d'éléments de E de norme 1 et tels que $\lim_{n \to +\infty} \|Ax_n - \lambda x_n\| = 0$.
On pose $m = \inf_{\|x\|=1} < Ax, x >$ et $M = \sup_{\|x\|=1} < Ax, x >$.
(d) Montrer que l'opérateur $A - mI$ est auto-adjoint, positif et que sa norme est égale à $M - m$. Montrer qu'il existe une suite (x_n) d'éléments de E, avec $\|x_n\| = 1$ tels que

$$\lim_{n \to +\infty} < (A - mI)x_n, x_n >= M - m$$

En écrivant $(A - MI)x_n = (A - mI)x_n - (M - m)x_n$, et en utilisant ce qui précède, montrer que $\lim_{n \to +\infty} \|Ax_n - Mx_n\|^2 = 0$ et en déduire que M appartient à $\sigma(A)$.
(e) En déduire que m appartient aussi à $\sigma(A)$ (on pourra considérer l'opérateur $-A$).

Chapitre 4

Opérateurs Compacts

Parmi tous les opérateurs continus dans un espace de Hilbert, on peut distinguer une classe importante d'opérateurs, dont les propriétés sont les plus proches de celles des opérateurs linéaires dans un espace de dimension finie. C'est la classe des opérateurs compacts, appelés encore opérateurs complètement continus.

4.1 Définitions et Propriétés

Nous commençons d'abord par rappeler la définition et quelques propriétés des ensembles compacts.

Définition 4.1.1. *Soit E un espace topologique séparé. Un sous-ensemble F est dit compact si de tout recrouvrement ouvert de F on peut extraire un sous recouvrement fini. Cela veut dire que, toute famille $\{V_j, j \in J\}$ d'ensembles ouverts dont la réunion contient F admet une sous-famille finie :*
$\{ V_{j(k)}, j(k) \in J, k = 1, 2 \ldots, n \}$ dont la réunion contient F.

La définition suivante est plus commode lorsque E est un espace métrique, cas dans lequel nous nous plaçons dans la suite.

Définition 4.1.2. *Soit E un espace métrique. Un sous-ensemble F de E est compact si toute suite d'éléments de F, contient une sous-suite convergente vers un élément de F.*

La propriété caractéristique suivante est souvent utile :

Proposition 4.1.3. *Soit E un espace métrique et F un sous-ensemble de E. Les propriétés suivantes sont équivalentes*

(a) *Le sous-ensemble F est compact*

(b) *Le sous-ensemble F est complet et pour tout $\epsilon > 0$, il existe un nombre fini d'éléments de F, $\{v_1, v_2, \ldots, v_n\}$, tels que*

$$F \subset \bigcup_{j=1}^{n} \mathcal{B}(v_j, \epsilon)$$

où $\mathcal{B}(v_j, \epsilon)$ est la boule ouverte de centre v_j et de rayon ϵ.

La propriété (b) exprime le fait que tout élément v de F est à une distance inférieure à ϵ d'au moins un élément parmi $\{v_1, v_2, \ldots, v_n\}$.

De ce qui précède, on peut voir que tout ensemble compact est un fermé, borné et complet. D'autre part, si E est de dimension finie, les sous-ensembles compacts de E sont exactement ceux qui sont fermés et bornés (théorème de Bolzano-Weierstrass), alors que dans un espace de dimension infinie, un sous-ensemble fermé et borné n'est pas nécessairement compact (voir encore la fin du chapitre I).

Définition 4.1.4. *Un sous-ensemble F de E est relativement compact si son adhérence $\overline{F}$ est compacte. Le sous-ensemble F est dit précompact si son complété est compact.*

Evidemment, lorsque E est lui-même complet, les deux notions sont équivalentes.

Nous sauterons la preuve de la proposition et de l'équivalence des définitions précédentes, et nous renvoyons le lecteur à un cours plus détaillé sur ces questions.

Définition 4.1.5. *Soit E un espace de Hilbert et A un endomorphisme de E. On dit que A est compact s'il transforme tout sous-ensemble borné de E en un ensemble relativement compact.*

En d'autres termes, A est un opérateur compact si, pour toute suite bornée (x_n) dans E, la suite (Ax_n) contient une sous-suite convergente.

On désignera par $\mathcal{K}(E)$ l'ensemble des opérateurs compacts de E dans lui-même.

Théorème 4.1.6. *Tout opérateur compact est borné, c'est-à-dire que l'on a l'inclusion $\mathcal{K}(E) \subset \mathcal{L}(E)$.*

Démonstration. Soit A un opérateur compact dans E. L'image de la boule unité de E est un ensemble relativement compact, donc borné. Il existe, alors, une constante $M > 0$, telle que

$$\|Ay\| \leq M, \quad \forall y \in E, \|y\| \leq 1$$

On en déduit que

$$\|Ax\| \leq M\|x\|, \quad \forall x \in E$$

L'opérateur A est donc continu et sa norme est majorée par M. □

Remarque 4.1.7. - La réciproque de ce théorème n'est pas vraie en général. Les opérateurs compacts sont très particuliers ; par exemple si E est de dimension infinie, l'identité n'est pas un opérateur compact, puisque cela entraînerait l'existence d'un voisinage de l'origine relativement compact dans E et par suite E serait de dimension finie.

En revanche, si E est de dimension finie, tout opérateur linéaire dans E est compact, puisque si F est un sous-ensemble borné de E, son image $A(F)$ est bornée et donc son adhérence $\overline{A(F)}$ est compacte.

Théorème 4.1.8. *L'ensemble $\mathcal{K}(E)$ est*

(a) *un sous-espace vectoriel fermé de $\mathcal{L}(E)$.*

(b) *un idéal bilatère de $\mathcal{L}(E)$, c'est-à-dire que si $A \in \mathcal{K}(E)$ et $B \in \mathcal{L}(E)$, alors AB et BA sont dans $\mathcal{K}(E)$.*

Démonstration. Il est facile de voir que $\mathcal{K}(E)$ est un sous-espace vectoriel de $\mathcal{L}(E)$; montrons qu'il est fermé. Soit (A_n) une suite d'opérateurs compacts qui converge vers un opérateur $A \in \mathcal{L}(E)$; il s'agit de prouver que A est un opérateur compact. Cela revient à prouver que $A(\mathcal{B}(0,1))$ est relativement compact, où $\mathcal{B}(0,1)$ est la boule unité de E. Soit $\epsilon > 0$, il existe un entier n_0 tel que $\|A_{n_0} - A\| < \epsilon/3$. Puisque A_{n_0} est compact, $A_{n_0}(\mathcal{B}(0,1))$ est relativement compact et il existe donc $x_1, x_2, \ldots, x_k$ dans $\mathcal{B}(0,1)$ tels que

$$A_{n_0}(\mathcal{B}(0,1)) \subset \bigcup_{j=1}^{k} \mathcal{B}(A_{n_0}x_j, \epsilon/3))$$

on en déduit que, pour tout $x \in \mathcal{B}(0,1)$,

$$\|Ax - Ax_j\| \leq \|Ax - A_{n_0}x\| + \|A_{n_0}x - A_{n_0}x_j\| + \|A_{n_0}x_j - Ax_j\| < \epsilon$$

c'est-à-dire que

$$A(\mathcal{B}(0,1)) \subset \bigcup_{j=1}^{k} \mathcal{B}(Ax_j, \epsilon)$$

ce qui termine la preuve de (a).

Soient, d'autre part, $A \in \mathcal{K}(E)$ et $B \in \mathcal{L}(E)$. Si une suite (x_j) est bornée dans E, il en sera de même de la suite (Bx_j) et, A étant compact, la suite (ABx_j) contient une sous-suite convergente. AB est donc un opérateur compact. On démontre de même que BA est un opérateur compact. □

Définition 4.1.9. *Un opérateur A est de rang fini si son image est de dimension finie. Le rang de A est la dimension de son image.*

L'ensemble des opérateurs de rang fini de E dans E sera noté $\mathcal{K}_0(E)$.

Exemple 4.1.10. - Soit E un espace de Hilbert séparable et soit (e_n) une base hilbertienne de E. Pour tout entier n, on désigne par P_n l'opérateur de projection orthogonale sur le sous-espace engendré par $\{ e_1, e_2, \ldots, e_n \}$. P_n est un opérateur de rang fini égal à n.

Exemple 4.1.11. - Sur l'espace de hilbert $L^2([0,\pi], dx)$, l'opérateur A défini par

$$Af(x) = \int_0^\pi \cos(x-s)f(s)\, ds, \quad \text{pour} \ \ f \in L^2([0,\pi], dx)$$

est manifestement un opérateur linéaire, continu et l'expression

$$Af(x) = \left(\int_0^\pi \cos(u)f(u)\, du\right)\cos x + \left(\int_0^\pi \sin(u)f(u)\, du\right)\sin x$$

montre que l'image de A est un sous-espace de dimension deux. L'opérateur A est donc de rang fini égal à 2.

Théorème 4.1.12. *Un opérateur borné A est de rang fini si, et seulement si, son adjoint A^* l'est aussi. Dans ce cas, A et A^* ont même rang.*

Démonstration. On suppose A de rang fini n. Soit $(e_i), i \leq n$, une base orthonormée de $\Im m(A)$ et soit P l'opérateur de projection orthogonale sur l'image de A. On a $A = PA$ et donc $A^* = A^*P$ (car P est auto-adjoint). On en déduit que pour tout $x \in E$

$$\begin{aligned} A^*x &= A^*\left(\sum_{i=1}^n \langle x, e_i\rangle e_i\right) \\ &= \sum_{i=1}^n \langle x, e_i\rangle A^* e_i \end{aligned}$$

L'image de A^* est donc engendrée par les vecteurs $(A^*e_i)_{i\leq n}$ et le rang de A^* est inférieur ou égal à celui de A. Comme $(A^*)^* = A$, on en déduit que A et son adjoint ont même rang. □

Théorème 4.1.13. *Tout opérateur de rang fini est compact.*

Démonstration. En effet, un opérateur de rang fini étant borné, il transforme un ensemble borné en un ensemble borné. Or, dans un espace de dimension finie, un ensemble borné est relativement compact. □

Il est facile de voir que $\mathcal{K}_0(E)$ constitue un sous-espace vectoriel de $\mathcal{K}(E)$.

Théorème 4.1.14. *Soit A un élément de $\mathcal{L}(E)$. Les assertions suivantes sont équivalentes*

(a) *L'opérateur A est compact.*
(b) *L'opérateur A^* est compact.*
(c) *L'opérateur A est limite d'une suite d'opérateurs de rang fini.*

Démonstration. L'assertion (c) implique l'assertion (a), car $\mathcal{K}_0(E)$ est inclus dans $\mathcal{K}(E)$ et celui-ci est fermé. Montrons que (a) implique (c) : Soit $\mathcal{B}(0,1)$ la boule unité de E, l'opérateur A étant compact, $A(\mathcal{B})(0,1)$ est relativement compact ou, ce qui revient au même, précompact. Il existe donc, pour tout $\epsilon > 0$, des éléments $y_1, y_2, \dots, y_k$ tels que

$$A(\mathcal{B}(0,1)) \subset \bigcup_{j=1}^{k} \mathcal{B}(y_j, \epsilon)$$

$\mathcal{B}(y_j, \epsilon)$ étant la boule de centre y_j et de rayon ϵ. Soi F le sous-espace vectoriel engendré par $y_1, y_2, \dots, y_k$ et P_F l'opérateur de projection orthogonale sur F. L'opérateur $P_F A$ est de rang fini et, pour tout x dans la boule $\mathcal{B}(0,1)$, Ax appartient à l'une des boules $\mathcal{B}(y_j, \epsilon)$, donc

$$\|Ax - P_F Ax\| = \mathrm{d}(Ax, F) = \inf_{y \in F} \|Ax - y\| \leq \epsilon$$

et par suite $\|A - P_F A\| \leq \epsilon$.
Enfin l'équivalence entre (a) et (b) découle du fait que l'adjoint d'un opérateur de rang fini est de rang fini et que la norme d'un opérateur continu est égale à la norme de son adjoint. □

Remarque 4.1.15. - Le caractère hilbertien de E est intervenu, de façon essentielle, par le projecteur orthogonal P_F. En fait, S. Banach a émis la conjecture que le théorème restait vrai lorsque E est un espace de Banach quelconque. Ce n'est qu'en 1972 que Per Enflo a montré que la réponse à cette conjecture est négative.

Les opérateurs compacts ont été introduits par Hilbert lors de l'étude des opérateurs intégraux. Ils les a appelé *opérateurs complètement continus* parce qu'ils possèdent une propriété de continuité spéciale que nous allons voir maintenant et qui, d'ailleurs, les caractérise (voir exercice 8).

Théorème 4.1.16. *Soit A un opérateur compact dans E. Alors, l'image par A de toute suite de E faiblement convergente est une suite (fortement) convergente.*

Démonstration. On raisonne par l'absurde. Soit (x_k) une suite faiblement convergente vers 0, et supposons que la suite (Ax_k) ne converge pas (en norme) vers 0. Il existe alors $\epsilon > 0$ et une sous-suite $(x_{k(n)})$ telle que $\|Ax_{k(n)}\| \geq \epsilon$, et ce pour tout n. La suite (x_k), étant faiblement convergente, est bornée. La suite $(Ax_{k(n)})$ admet donc une sous-suite qui converge

(en norme), soit y sa limite. Pour simplifier les notations, on peut supposer que

$$\lim_{n\to+\infty} \|Ax_{k(n)} - y\| = 0$$

Or, la suite $(Ax_{k(n)})$ converge faiblement vers 0, donc nécessairement $y = 0$, c'est-à-dire

$$\lim_{n\to+\infty} \|Ax_{k(n)}\| = 0$$

ce qui contredit l'hypothèse. □

Corollaire 4.1.17. *Soit (e_k), $k \in \mathbb{N}$, une suite orthonormée dans E. Si A un élément de $\mathcal{K}(E)$, alors $\lim_{k\to+\infty} \|Ae_k\| = 0$.*

Démonstration. Pour tout x dans E, la série $\sum_k |\langle x, e_k\rangle|^2$ est convergente et son terme général $\langle x, e_k\rangle$ tend vers 0 lorsque k tend vers l'infini. Cela traduit le fait que la suite (e_k) est faiblement convergente vers 0, le théorème précédent permet de conclure. □

Un exemple important d'opérateurs compacts est donné par le théorème qui suit.

Théorème 4.1.18. *Soient E un espace de Hilbert séparable et (e_n) une base hilbertienne de E. Soit $\lambda = (\lambda_n)$ une suite bornée dans $\mathbb{C}$ alors l'opérateur de multiplication par λ défini par*

$$A_\lambda e_n = \lambda_n e_n$$

est compact si, et seulement si, λ_n tend vers zéro quand n tend vers l'infini.

Démonstration. Soit P_n l'opérateur de projection orthogonale sur le sous-espace engendré par $\{ e_1, e_2, \ldots, e_n \}$. L'opérateur P_nA_λ est un opérateur de rang fini et on a

$$(A_\lambda - P_nA_\lambda)e_j = \begin{cases} \lambda_j e_j, & \text{si } j \geq n+1\,; \\ 0, & \text{si } j \leq n. \end{cases}$$

et par suite $\|A_\lambda - P_nA_\lambda\| = \sup_{j>n} |\lambda_j|$. Si la suite (λ_n) tend vers zéro quand n tend vers l'infini, on en déduit que l'opérateur A_λ est limite d'une suite d'opérateurs de rang fini, donc compact (d'après le théorème 1.14). Inversement, si A_λ est compact, le corollaire 1.17 assure que

$$\lim_{n\to\infty} \|A_\lambda e_n\| = \lim_{n\to\infty} |\lambda_n| = 0$$

ce qui termine la preuve. □

Un deuxième exemple d'opérateurs compacts est donné par la famille des opérateur de Hilbert-Schmidt qui ne sont définis que si l'espace de Hilbert est séparable (ce que nous supposerons dans la suite). Pour les introduire, nous avons besoin du résultat qui suit.

Proposition 4.1.19. *On suppose E séparable et on désigne par (e_n) une base hilbertienne de E. Pour tout élément A de $\mathcal{L}(E)$, le nombre (fini ou infini)*

$$|||A|||^2 = \sum_{n=1}^{\infty} \|Ae_n\|^2$$

ne dépend pas de la base hilbertienne considérée et on a

$$|||A||| = |||A^*|||$$

Démonstration. Soit (f_n) une autre base hilbertienne de E. On a

$$\begin{aligned}\sum_{p=1}^{\infty} \|Ae_p\|^2 &= \sum_{p=1}^{\infty}\sum_{q=1}^{\infty} |\langle Ae_p, f_q\rangle|^2 = \sum_{p=1}^{\infty}\sum_{q=1}^{\infty} |\langle e_p, A^*f_q\rangle|^2 \\ &= \sum_{q=1}^{\infty}\sum_{p=1}^{\infty} |\langle e_p, A^*f_q\rangle|^2 = \sum_{q=1}^{\infty} \|A^*f_q\|^2\end{aligned}$$

En particulier, en prenant la même base, on obtient

$$\sum_{p=1}^{\infty} \|Ae_p\|^2 = \sum_{p=1}^{\infty} \|A^*e_p\|^2$$

ce qui montre que $|||A||| = |||A^*|||$. On en déduit ensuite que $|||A|||$ ne dépend pas de la base considérée, puisque

$$\sum_{p=1}^{\infty} \|Ae_p\|^2 = \sum_{p=1}^{\infty} \|A^*f_p\|^2 = \sum_{p=1}^{\infty} \|Af_p\|^2 \qquad \square$$

Définition 4.1.20. *Un opérateur de Hilbert-Schmidt sur E est un élément A de $\mathcal{L}(E)$ pour lequel $|||A|||$ est fini*

$$|||A||| = \Big(\sum_{n=1}^{\infty} \|Ae_n\|^2\Big)^{\frac{1}{2}} < \infty$$

Le nombre $|||A|||$ s'appelle la norme de Hilbert-Schmidt de A.

Notons que tout opérateur de Hilbert-Schmidt est borné et on a l'inégalité $\|A\| \leq |||A|||$ car, pour tout $x \in E$,

$$\|Ax\|^2 = \sum_{i=1}^{\infty} |\langle Ax, e_i\rangle|^2 = \sum_{i=1}^{\infty} |\langle x, A^*e_i\rangle|^2 \leq \|x\|^2|||A|||^2$$

Théorème 4.1.21. *Tout opérateur de Hilbert-Schmidt est compact.*

Démonstration. Soit (e_n) une base hilbertienne de E et soit A un opérateur de Hilbert-Schmidt sur E. Pour tout n, on désigne par P_n l'opérateur de projection orthogonale sur le sous-espace engendré par $\{e_1, e_2, \ldots, e_n\}$. L'opérateur $A_n = AP_n$ est de rang fini égal à n et on a $A_n e_j = Ae_j$, si $j \leq n$ et $A_n e_j = 0$, si $j > n+1$, donc

$$(A - A_n)e_j = \begin{cases} 0, & \text{si } j \leq n\,; \\ Ae_j, & \text{si } j \geq n+1. \end{cases}$$

Il en résulte que

$$\|A - A_n\|^2 \leq |||A - A_n|||^2 = \sum_{j=n+1}^{\infty} \|Ae_j\|^2$$

Le second membre de cette inégalité tend vers zéro lorsque n tend vers l'infini, puisque c'est le reste d'une série convergente. L'opérateur A est donc limite d'une suite d'opérateurs de rang fini. □

Lorsque E est l'espace $L^2(X, \Omega, \mu)$, on a une description précise des opérateurs de Hilbert-Schmidt. Plus exactement, soit (X, Ω, μ) un espace mesuré σ-fini et supposons que l'espace de Hilbert $L^2(X, \Omega, \mu)$ soit séparable, on a :

Proposition 4.1.22. *Soit $k \in L^2(X \times X, \Omega \times \Omega, \mu \otimes \mu)$. Pour tout élément f de $L^2(X, \Omega, \mu)$, la fonction*

$$Kf(x) = \int k(x, y) f(y) d\mu(y)$$

est définie pour presque tout x et appartient à $L^2(X, \Omega, \mu)$. L'application qui à f associe Kf est un opérateur de Hilbert-Schmidt sur E et on a

$$|||K|||^2 = \int\int |k(x, y)|^2 d\mu(x) d\mu(y)$$

Inversement, tout opérateur de Hilbert-Schmidt sur $L^2(X, \Omega, \mu)$ est du type décrit ci-dessus.

Le lemme suivant est utile pour démontrer la proposition. On en donne une généralisation en exercice.

Lemme 4.1.23. *Soit (ϕ_n) une base hilbertienne de $L^2(X, \Omega, \mu)$. La famille (Φ_{pq}), définie sur $X \times X$ par $\Phi_{pq}(x, y) = \phi_p(x)\overline{\phi_q(y)}$, est une base hilbertienne de $L^2(X \times X, \Omega \times \Omega, \mu \otimes \mu)$.*

Démonstration. Puisque pour (p, q) et (i, j) dans $\mathbb{N}^2$

$$\int \Phi_{pq}\overline{\Phi_{ij}}d\mu \otimes \mu = \int\int \phi_p(x)\overline{\phi_q}(y)\overline{\phi_i}(x)\phi_j(y)d\mu(x)d\mu(y)$$
$$= \langle \phi_p, \phi_i \rangle \langle \phi_j, \phi_q \rangle$$

la famille (Φ_{pq}) est orthonormée. Soit F une fonction de carré intégrable sur l'espace produit $X \times X$. Pour presque tout $y \in X$,

$$\int |F(x, y)|^2 d\mu(x) < \infty$$

Il en résulte que la fonction

$$F_i(y) = \int \overline{F}(x, y)\phi_i(x)d\mu(x)$$

est définie presque partout, appartient à $L^2(X, \Omega, \mu)$ et on a

$$\|F_i\|^2 = \sum_j |\langle \phi_j, F_i \rangle|^2 = \sum_j |\int \overline{F_i}(y)\phi_j(y)d\mu(y)|^2$$
$$= \sum_j |\int\int F(x, y)\overline{\phi_i}(x)\phi_j(y)d\mu(x)d\mu(y)|^2$$
$$= \sum_j |\langle F, \Phi_{ij} \rangle|^2$$

Ainsi, si F est orthogonale à Φ_{ij} pour tout couple (i, j), alors $F_i = 0$, quel que soit l'entier i, donc $x \mapsto F(x, y)$ est nulle pour presque tout y, c'est-à-dire que $F = 0$. (Φ_{pq}) est donc une base hilbertienne. □

Démonstration la proposition : D'après le théorème de Fubini, pour presque tout x, l'application $y \mapsto |k(x, y)|^2$ est intégrable et l'inégalité de Cauchy-Schwarz donne

$$|Kf(x)|^2 \leq \|f\|^2 \int |k(x, y)|^2 d\mu(y)$$

Si Kf est mesurable, on aura

$$\int |Kf(x)|^2 d\mu(x) \leq \|f\|^2 \int\int |k(x, y)|^2 d\mu(x)d\mu(y)$$

La fonction Kf est donc dans $L^2(X, \Omega, \mu)$ et $\|Kf\| \leq \|k\| \, \|f\|$. Cela montre que K est un opérateur borné et $\|K\| \leq \|k\|$.

Montrons que la fonction Kf est mesurable. Soit Δ un ensemble intégrable et χ_Δ sa fonction caractéristique. L'application qui à (x, y) associe

$\chi_\Delta(x)f(y)$ est de carré intégrable sur $X \times X$; comme k est aussi de carré intégrable sur $X \times X$, on en déduit que l'application qui à (x,y) associe $k(x,y)\chi_\Delta(x)f(y)$ est intégrable. D'après le théorème de Fubini, pour presque tout x, l'application $y \mapsto k(x,y)\chi_\Delta(x)f(y)$ est intégrable et l'application

$$x \mapsto \int k(x,y)\chi_\Delta(x)f(y)d\mu(y)$$

est intégrable. Cela montre que, pour tout ensemble intégrable Δ, la fonction $\chi_\Delta Kf$ est mesurable. L'espace $L^2(X,\Omega,\mu)$ étant supposé σ-fini, il en résulte que Kf est mesurable.

Montrons que K est un opérateur de Hilbert-Schmidt. Soit (ϕ_n) une base hilbertienne de $L^2(X,\Omega,\mu)$, on a

$$\begin{aligned}\langle K\phi_q, \phi_p\rangle &= \iint k(x,y)\phi_q(y)\overline{\phi_p}(x)d\mu(x)d\mu(y)\\ &= \iint k(x,y)\overline{\overline{\phi_q}(y)\phi_p(x)}d\mu(x)d\mu(y)\end{aligned}$$

c'est-à-dire $\langle K\phi_q, \phi_p\rangle = \langle k, \Phi_{pq}\rangle$. Le lemme précédent permet alors d'écrire

$$\sum_{p,q=1}^{\infty} |\langle K\phi_p, \phi_q\rangle|^2 = \sum_{p,q=1} |\langle k, \Phi_{pq}\rangle|^2 = \iint |k(x,y)|^2 d\mu \otimes \mu$$

On en déduit l'égalité $|||K||| = \|k\|$.
Inversement, soit A un opérateur de Hilbert-Schmidt sur $L^2(X,\Omega,\mu)$ et (ϕ_n) une base hilbertienne. Posons

$$a_{ij} = \langle A\phi_j, \phi_i\rangle \quad \text{et} \quad k(x,y) = \sum_{i,j} a_{ij}\phi_i(x)\overline{\phi_j}(y)$$

Cette dernière série converge dans $L^2(X \times X, \Omega \times \Omega, \mu \otimes \mu)$, car $(\phi_i \otimes \overline{\phi_j})$ est une base hilbertienne et $\sum_{i,j} |a_{ij}|^2 < \infty$; il en résulte que k est de carré intégrable relativement à la mesure $\mu \otimes \mu$. D'autre part, soit f et g dans $L^2(X,\Omega,\mu)$ et posons

$$g = \sum_i z_i\phi_i, \quad Af = \sum_i y_i\phi_i \quad \text{avec} \quad y_i = \sum_j a_{ij}x_j$$

On a

$$\begin{aligned}\langle Af, g\rangle &= \sum_i y_i\overline{z_i} = \sum_{ij} a_{ij}x_j\overline{z_i}\\ &= \langle k, g \otimes \overline{f}\rangle = \iint k(x,y)f(y)\overline{g}(x)d\mu(y)d\mu(x)\\ &= \int \Big(\int k(x,y)f(y)\, d\mu(y)\Big)\overline{g}(x)\, d\mu(x)\end{aligned}$$

Il en résulte que A est l'opérateur intégral associé au noyau k. □

Proposition 4.1.24. *L'ensemble des opérateurs de Hilbert-Schmidt est un idéal bilatère de $\mathcal{L}(E)$ et si A est un opérateur de Hilbert Schmidt, pour toute base hilbertienne (e_n) de E, on a*

$$|||A|||^2 = \sum_{i,j=1}^{\infty} |a_{ij}|^2 \quad \text{où} \quad a_{ij} = \langle Ae_j, e_i \rangle$$

Démonstration. Si A est un opérateur de Hilbert-Schmidt et S un élément de $\mathcal{L}(E)$, on a

$$\sum_{j=1}^{\infty} \|SAe_j\|^2 \leq \|S\|^2 \sum_{j=1}^{\infty} \|Ae_j\|^2 = \|S\|^2 |||A|||^2$$

L'opérateur SA est donc un opérateur de Hilbert-Schmidt. De même, S^*A^* est un opérateur de Hilbert-Schmidt et il en sera de même de son adjoint AS, ce qui prouve (a). L'assertion (b) résulte de l'égalité

$$\|Ae_j\|^2 = \sum_{i=1}^{\infty} |\langle Ae_j, e_i \rangle|^2$$

L'assertion (3) est immédiate. □

EXERCICES

1. Soit $A \in \mathcal{L}(E)$. Montrer que si A^*A est un opérateur compact alors A est un opérateur compact. En déduire que A est compact si et seulement si A^* est compact. (c'est une autre démonstration du théorème 1.14).

Solution : 1) Soit (x_n) une suite bornée de E. Puisque A^*A est compact, on peut extraire de (x_n) une sous-suite (x'_n) telle que $(A^*Ax'_n)$ soit convergente, on désigne par y sa limite : $\lim_{n\to\infty} \|A^*Ax'_n - y\| = 0$. On peut extraire de (x'_n) une sous-suite (x''_n) faiblement convergente, on désigne pat x la limite faible de (x''_n). On a

$$\|Ax''_n\|^2 = \langle Ax''_n, Ax''_n \rangle = \langle A^*Ax''_n, x''_n \rangle$$

Il vient

$$\begin{aligned} \left| \|Ax''_n\|^2 - \langle y, x \rangle \right| &\leq \left| \langle A^*Ax''_n, x''_n \rangle - \langle y, x''_n \rangle \right| + |\langle y, x''_n \rangle - \langle y, x \rangle| \\ &\leq \|A^*Ax''_n - y\| \|x''_n\| + |\langle y, x''_n \rangle - \langle y, x \rangle| \end{aligned}$$

Comme (x''_n) est bornée et $A^*Ax''_n$ converge vers y, le premier terme du second membre tend vers 0 quand n tend vers l'infini. Il en est de même du second terme car (x''_n) converge faiblement vers x. On a

donc montré que la suite (Ax''_n) est convergente et par suite A est un opérateur compact.
2) Si maintenant A est compact, il en sera de même de AA^*, car le produit d'un opérateur compact par un opérateur borné est compact ; l'étape 1) montre alors que A^* est compact. On montre de même que si A^* est compact, A l'est aussi. □

2. Montrer qu'un opérateur $A \in \mathcal{L}(E)$ est compact si, et seulement si, l'application $S \mapsto SAS$, de $\mathcal{L}(E)$ dans lui-même est compacte.

3. Soit $\phi : [a, b] \to \mathbb{R}$ une fonction continue. Soit A_ϕ l'opérateur de multiplication par ϕ, défini sur $L^2((a, b), dx)$ par

$$A_\phi f = \phi f$$

Montrer que A_ϕ est continu et que s'il existe $x \in [a, b]$ tel que $\phi(x) \neq 0$, alors A_ϕ n'est pas compact.

Solution : D'abord il est clair que A_ϕ est un opérateur continu. Supposons que ϕ ne s'annule pas en un point de $[a, b]$, comme elle est continue il existe un intervalle $[c, d]$ inclus dans $[a, b]$ dans lequel ϕ ne s'annule pas. Si A_ϕ est compact dans $L^2[a, b]$, sa restriction à $L^2[c, d]$ est à fortiori un opérateur compact. Mais cette restriction est un opérateur inversible (son inverse est l'opérateur de multiplication par $1/\phi(x)$, $c \leq x \leq d$), il en résulte que l'identité est un opérateur compact dans $L^2[c, d]$, ce qui est impossible. □

4. Montrer que dans un espace normé de dimension infinie, si un opérateur compact est inversible, alors son inverse n'est pas borné.

Solution : Si A est compact et inversible, A^{-1} ne peut pas être borné sinon, AA^{-1} serait compact, ce qui exige que E soit de dimension finie. □

5. Un opérateur compact dans un espace de Hilbert (de dimension infinie) peut-il être solution d'une équation algébrique $\sum_{k=0}^{n} c_k A^k = 0$ (en convenant que $A^0 = I$) ?

Solution : Si $c_0 = 0$, la réponse est *oui.* A titre d'exemple, on peut considérer l'opérateur de projection orthogonal P sur un espace de dimension finie, c'est un opérateur compact et il vérifie l'équation $P - P^2 = 0$. En revanche, si c_0 est non nul, la réponse est *non.* En effet, dans ce cas l'équation peut s'écrire sous la forme

$$b_1 A + b_2 A^2 + \cdots + b_n A^n = I \quad \text{avec} \quad b_k = c_k / c_0$$

et l'opérateur A serait compact et inversible dans $\mathcal{L}(E)$, ce qui contredit le fait que la dimension de E est infinie. □

6. Soit E un espace de Hilbert et $A \in \mathcal{L}(E)$. Montrer que si A transforme une suite faiblement convergente en une suite fortement convergente, alors il est compact (c'est la réciproque du théorème 1.16).

Solution : Comme la boule unité est faiblement compacte, voir le théorème 3.11 du Chapitre I, l'hypothèse dit que l'opérateur A transforme la boule unité en un ensemble relativement compact, ce qui est la définition de la compacité de A. □

7. Montrer que si $p < q$, l'injection de l'espace de Sobolev $H^q[0, 2\pi]$ dans l'espace de Sobolev $H^p[0, 2\pi]$ est compacte. Ces espaces ont été introduits au chapitre I, exemple 1.14.

   Solution : On désigne par I l'injection de $H^q[0, 2\pi]$ dans $H^p[0, 2\pi]$ et par I_n l'opérateur de projection orthogonale sur le sous-espace engendré par $\{ e_m; |m| \leq n \}$. La projection d'un élément f de $H^q[0, 2\pi]$ s'écrit $I_n(f) = \sum_{|m| \leq n} c_m(f) e_m$. On vérifie alors, que

$$\begin{aligned} \|(I - I_n)f\|_p^2 &= \sum_{|m| \geq n+1} (1 + m^2)^p |c_m(f)|^2 \\ &\leq \frac{1}{(1 + n^2)^{q-p}} \|f\|_q^2 \end{aligned}$$

   Ainsi, l'identité l'injection de $H^q[0, 2\pi]$ dans $H^p[0, 2\pi]$ est limite d'une suite d'opérateurs de rang fini, elle est donc compacte. □

8. Soit $[a, b]$ un intervalle fermé borné de $\mathbb{R}$. On désigne par $C^k[a, b]$ l'espace des fonctions continûment dérivables sur $[a, b]$ jusu'à l'ordre k. On munit cet espace de la norme suivante : $\|f\|_{(k)} = \sup_{j \leq k} \|f^{(j)}\|_\infty$. La convergence dans $C^k[a, b]$ équivaut à la convergence uniforme des fonctions avec leurs dérivées jusqu'à l'ordre k.
   (a) Montrer que $C^k[a, b]$, muni de la norme ci-dessus est un espace de Banach.
   (b) Montrer que l'injection de l'espace $C^{k+1}[a, b]$ dans l'espace $C^k[a, b]$ est compacte.

   Solution : Soit $\mathcal{F}$ une famille bornée d'éléments de $C^1[a, b]$. Il existe une constante M telle que $\|f\|_{(1)} \leq M$, pour tout $f \in \mathcal{F}$. La famille $\mathcal{F}$ est évidemment bornée dans $C[a, b]$, de plus elle est équicontinue, car

$$|f(x + h) - f(x)| = |\int_x^{x+h} f'(t)\, dt| \leq M|h|, \quad \forall f \in \mathcal{F}$$

   Le théorème d'Ascoli-Arzela (voir l'annexe) permet d'en déduire que $\mathcal{F}$ est relativement compact. Ainsi, on vient de montrer que l'injection de $C^1[a, b]$ dans $C[a, b]$ est compacte. On montre de même que l'injection de l'espace $C^{k+1}[a, b]$ dans l'espace $C^k[a, b]$ est compacte. □

9. Soit H le sous-espace de $L^2(\mathbb{T})$, des fonctions f dont les coefficients de Fourier vérifient $c_m(f) = 0$, $\forall m < 0$. Il peut être considéré comme l'image de l'opérateur de projection orthogonale sur le sous-espace de

$L^2(\mathbb{T})$, engendré par $(e_m)_{m\geq 0}$,

$$Pf(e^{i\theta}) = \sum_{m=0}^{\infty} c_m(f)e^{im\theta}, \quad f \in L^2(\mathbb{T})$$

Soit $\phi \in C(\mathbb{T})$ et définissons l'opérateur de Toeplitz $T_\phi : H \to H$, par $T_\phi f = P(\phi f)$.
(a) Montrer que $\|T_\phi\| \leq \|\phi\|_\infty$
(b) Par un calcul explicite, montrer que pour $\phi = e_k$, l'opérateur $T_{e_k}T_{e_j} - T_{e_k e_j}$ est compact sur H.
(c) Montrer que, pour ϕ et ψ dans $C(\mathbb{T})$, l'opérateur $T_\phi T_\psi - T_{\phi\psi}$ est compact sur H. (on approximera ϕ et ψ par des combinaisons linéaires finies des (e_k).)

10. Montrer que l'image de tout opérateur compact est séparable.

11. Soit E un espace de Hilbert et A un élément de $\mathcal{L}(E)$. Montrer que A est compact si, et seulement si, chaque fois que deux suites (x_n) et (y_n) convergent faiblement vers x et y respectivement, alors la suite $(\langle Ax_n, y_n\rangle)$ converge vers $\langle Ax, y\rangle$. (cet exercice est à rapprocher de l'exercice 3, section 1, chapitre III).

Solution : Les hypothèses traduisent le fait que le graphe de l'opérateur A est fermé. La solution est donc une conséquence immédiate du théorème du graphe fermé. □

12. Soient (X_i, Ω_i, μ_i), pour $1 \leq i \leq 2$, deux espaces mesurés et σ-finis. Soient (ϕ_n) une base hilbertienne de $L^2(X_1, \Omega_1, \mu_1)$ et (ψ_n) une base hilbertienne de $L^2(X_2, \Omega_2, \mu_2)$. En suivant la preuve du lemme 1.23, montrer que $\{\, \phi_p \otimes \psi_q; p, q \in \mathbb{N} \,\}$ est une base hilbertienne de l'espace $L^2(X_1 \times X_2, \Omega \times \Omega_2, \mu_1 \otimes \mu_2)$.

4.2 Spectre d'un opérateur compact

Soient E un espace de Hilbert et A un opérateur compact dans E. On sait que si E n'est pas de dimension finie, $\lambda = 0$ est, toujours, dans le spectre de A ; sinon, l'opérateur $I = AA^{-1}$ serait compact ce qui impliquerait que E est de dimension finie (Remarque 1.7). Dans la suite donc, λ désignera toujours, sauf mention du contraire, un nombre complexe non nul.

Proposition 4.2.1. *Soit $A \in \mathcal{K}(E)$. Si $A - \lambda I$ est surjectif, alors il est injectif.*

Démonstration. Supposons qu'il existe $x_1 \neq 0$ tel que $Ax_1 = \lambda x_1$. Posons $B = A - \lambda I$. On a $Bx_1 = 0$, on a aussi

$$\ker(B) \subset \ker(B^2) \subset \cdots \subset \ker(B^n) \subset \cdots$$

et chacun de ces sous-espaces est fermé. Nous allons montrer que ces inclusions sont, toutes, strictes. En effet, puisque B est surjectif, il existe x_2 tel que $Bx_2 = x_1$, il existe x_3 tel que $Bx_3 = x_2$ etc. Par induction, on construit une suite (x_n) telle que $Bx_{n+1} = x_n$ et $x_n \neq 0$ puisque $x_1 \neq 0$. D'autre part, $x_n \in \ker(B^n)$, car

$$B^n x_n = B^{n-1}(Bx_n) = B^{n-1}(x_{n-1}) = \cdots = Bx_1 = 0$$

Mais $x_n \notin \ker(B^{n-1})$, car $B^{n-1}(x_n) = \cdots = B(x_2) = x_1 \neq 0$, donc $\ker(B^{n-1})$ est strictement inclus dans $\ker(B^n)$. Choisissons, pour tout $n \geq 1$, un vecteur unitaire e_n appartenant à $\ker(B^n)$ et orthogonal à $\ker(B^{n-1})$. Comme Be_n appartient à $\ker(B^{n-1})$, on a

$$\|Ae_n\|^2 = \|Be_n + \lambda e_n\|^2 = \|Be_n\|^2 + |\lambda|^2 \geq |\lambda|^2$$

Ainsi, (e_n), $n \geq 1$, est une suite orthonormée, dont l'image (Ae_n), $n \in \mathbb{N}$, ne converge pas vers 0, ce qui est impossible d'après le corollaire 1.17. □

Pour la réciproque de cette proposition, on a besoin du lemme suivant

Lemme 4.2.2. *Soit $A \in \mathcal{K}(E)$. Si $A - \lambda I$ est injectif, alors son image est fermée dans E.*

Démonstration. Soient $y \in \overline{\Im m(A - \lambda I)}$ et (y_n) une suite de $\Im m(A - \lambda I)$ qui converge vers y. On pose

$$y_n = Ax_n - \lambda x_n$$

• Si (x_n) contient une sous-suite bornée, alors, A étant compact, (x_n) contient aussi une sous-suite (x_{n_k}) telle que (Ax_{n_k}) converge. Comme

$$x_{n_k} = \frac{Ax_{n_k} - y_{n_k}}{\lambda}$$

la suite (x_{n_k}) converge vers un élément x qui vérifie $Ax - \lambda x = y$.
• Si (x_n) ne contient aucune sous-suite bornée, la suite $\|x_n\|$ tend vers l'infini avec n. Posons $z_n = \|x_n\|^{-1} x_n$, il vient

$$\|z_n\| = 1 \quad \text{et} \quad \lim_{n\to\infty} (A - \lambda I) z_n = 0$$

Comme A est compact, (z_n) contient une sous-suite (z_{n_k}) telle que (Az_{n_k}) converge. On en déduit que la suite (z_{n_k}) est convergente et si z est sa limite, on aura $\|z\| = 1$ et $Az - \lambda z = 0$. Ce qui contredit l'hypothèse $A - \lambda I$ est injectif. □

Lemme 4.2.3. *Soit $A \in \mathcal{K}(E)$. Si l'opérateur $A - \lambda I$ est injectif, alors il est surjectif.*

Démonstration. Puisque $B = A - \lambda I$ est injectif, son image E_1 est fermée d'après le lemme précédent. Posons, pour tout $n \geq 1$, $E_n = B^n(E)$. On a là une suite décroissante de sous-espaces fermés

$$\cdots \subset E_n \subset E_{n-1} \subset \cdots \subset E_2 \subset E_1$$

Si l'on suppose que $E_1 \neq E$, ces inclusions seront, en fait, strictes. En effet, si x_0 est un élément non nul de E qui n'appartient pas à E_1, il est facile de voir que pour tout $n \geq 1$, $B^n x_0 \in E_n$ et $B^n x_0 \notin E_{n+1}$. Choisissons donc $e_n \in E_n$ de norme 1 et orthogonal à E_{n+1}. Comme $Be_n \in E_{n+1}$:

$$\|Ae_n\|^2 = \|Be_n\|^2 + |\lambda|^2 \geq |\lambda|^2$$

ce qui contredit le fait que (Ae_n) doit tendre vers 0. □

Les résultats précédents se résument comme suit

Théorème 4.2.4. *Soit A un opérateur compact dans un espace de Hilbert E. Pour tout λ complexe non nul, l'opérateur $A - \lambda I$ est injectif si, et seulement si, il est surjectif. Autrement dit*

$$\sigma(A) \setminus \{0\} = \sigma_p(A) \setminus \{0\}$$

Corollaire 4.2.5. (Alternative de Fredholm) *Soit A un élément de $\mathcal{K}(E)$ et μ un nombre complexe. Alors, deux possibilités peuvent se présenter pour l'équation*

$$x - \mu Ax = y$$

- Ou bien, quel que soit le choix du second membre y, elle possède une solution unique x.
- Ou bien l'équation homogène $x - \mu Ax = 0$ admet une solution non nulle.

Démonstration. Il est évident que pour $\mu = 0$, c'est la première possibilité qui se réalise. Soit donc $\mu \neq 0$ et posons $\lambda = \mu^{-1}$. L'équation $x - \mu Ax = y$ est alors équivalente à l'équation

$$(A - \lambda I)x = -\lambda y$$

Si λ n'est pas dans le spectre de A, alors $A - \lambda I$ est inversible, et c'est le premier cas qui a lieu ; si λ appartient au spectre de A, le théorème précédent montre que λ est, nécessairement, une valeur propre de A (car $\lambda \neq 0$) et c'est alors le deuxième cas qui se présente. □

Théorème 4.2.6. *Si A un opérateur compact alors, pour tout complexe λ non nul, le noyau de $A - \lambda I$ est de dimension finie.*

Démonstration. Il suffit de montrer que la boule unité du sous-espace $\ker(A-\lambda I)$ est relativement compacte. Soit (x_n) une suite dans $\ker(A-\lambda I)$, de norme inférieure ou égale à 1. On a

$$x_n = \frac{Ax_n}{\lambda}$$

L'opérateur A étant compact, (Ax_n) contient une sous-suite convergente, et la relation précédente montre que la suite (x_n), elle-même, contient une sous-suite convergente. □

Corollaire 4.2.7. *Pour tout entier $n \geq 1$, $\ker\big((A-\lambda I)^n\big)$ est de dimension finie.*

Démonstration. En effet, on peut écrire

$$\begin{aligned}(A-\lambda I)^n &= A^n - \binom{n}{1}\lambda A^{n-1} + \binom{n}{2}\lambda^2 A^{n-2} \cdots + (-1)^n \lambda^n I \\ &= AB + (-1)^n \lambda^n I\end{aligned}$$

où B est un opérateur borné, et le théorème précédent permet alors de conclure. □

Théorème 4.2.8. *Soit A un opérateur compact. Si (λ_n) est une suite de valeurs propres de A alors ou bien cette suite est finie, ou bien elle tend vers 0 quand n tend vers l'infini.*

Démonstration. Supposons que la suite (λ_n) ne tende pas vers 0. Il existe alors $\epsilon > 0$ et une sous-suite (λ_{n_k}) satisfaisant $|\lambda_{n_k}| \geq \epsilon$. Pour simplifier les notations, on peut supposer que $|\lambda_n| \geq \epsilon$. Soit, pour tout $k \geq 1$, x_k un vecteur propre associé à λ_k, et soit N_k le sous-espace engendré par $\{x_1, x_2, \ldots, x_k\}$.
La suite (N_k) est strictement croissante. Pour le voir, Il suffit de montrer que, pour tout k, $\{x_1, \ldots, x_k\}$ est une famille libre. Supposons donc que $\{x_1, \ldots, x_{k-1}\}$ soit libre et que $x_k = \sum_1^{k-1} \alpha_i x_i$. Il vient

$$0 = (A - \lambda_k I)x_k = \sum_{i=1}^{k-1} \alpha_i (A - \lambda_k I)x_i = \sum_{i=1}^{k-1} \alpha_i (\lambda_i - \lambda_k) x_i$$

Comme $\lambda_k - \lambda_i \neq 0$ pour tout $i \leq k-1$, on a nécessairement $\alpha_i = 0$, $\forall i \leq k-1$ et donc $x_k = 0$, ce qui est absurde. On peut donc choisir, pour tout $k \geq 1$, un élément e_k de N_k, de norme 1 et orthogonal à N_{k-1}. Les vecteurs e_k et $(A-\lambda_k I)e_k$ sont alors orthogonaux, car $(A-\lambda_k I)N_k \subset N_{k-1}$. On en déduit que

$$\begin{aligned}\|Ae_k\|^2 &= \|(A-\lambda_k I)e_k + \lambda_k e_k\|^2 \\ &= \|(A-\lambda_k I)e_k\|^2 + |\lambda_k|^2 \geq |\lambda_k|^2\end{aligned}$$

Mais cela contredit le fait que $\lim_{k\to\infty} \|Ae_k\| = 0$. □

On a maintenant une description précise du spectre d'un opérateur compact dans un espace de Hilbert.

Théorème 4.2.9. *Soit A un opérateur compact dans un espace de Hilbert E, alors*

(a) *Le spectre de A est au plus dénombrable. Chaque point du spectre est isolé, à l'exception possible de 0.*

(b) *Si l'espace E n'est pas de dimension finie alors 0 appartient au spectre de A .*

(c) *Tout complexe non nul du spectre de A est une valeur propre et l'espace propre correspondant est de dimension finie.*

Démonstration. Puisque $\sigma(A)$ est compact, il suffit de montrer que tout $\lambda \in \sigma(A) \setminus \{0\}$ est isolé. Or, si un tel λ n'est pas isolé, il existerait une suite (λ_k), d'éléments dans $\sigma(A)$ non nuls et distincts deux à deux, qui convergerait vers λ. Comme $\sigma(A) \setminus \{0\} = \sigma_p(A) \setminus \{0\}$, les λ_k sont des valeurs propres et le théorème 2.8 dit que nécessairement (λ_k) converge vers 0, ce qui contredit l'hypothèse $\lambda \neq 0$. □

Remarque 4.2.10. - Soient E un espace de Hilbert séparable, $(e_n)_{n\geq 1}$ une base hilbertienne et $\lambda = (\lambda_n)_{n\geq 1}$ une suite de nombres complexes qui tend vers zéro. L'opérateur A_λ de multiplication par la suite λ qui, à $x = (x_n) \in E$ associe l'élément $A_\lambda x = (\lambda_n x_n)$ est compact (théorème 1.18). Si l'on suppose que $\lambda_n \neq 0, \forall n$, alors 0, bien qu'il soit dans $\sigma(A_\lambda)$, n'est pas une valeur propre. En revanche, si l'on prend $\lambda_1 = 0$ et $\lambda_n \neq 0, \forall n \geq 2$, alors 0 serait une valeur propre avec e_1 comme vecteur propre associé . Notons aussi que, même si $\lambda = 0$ est une valeur propre, le sous-espace propre associé, c'est-à-dire $\ker(A)$, n'est pas en général de dimension finie. C'est le cas, par exemple, de l'opérateur de multiplication dans $\ell^2(\mathbb{N})$ par une suite $\lambda = (\lambda_n)$ qui est nulle à partir d'un certain rang (ou simplement la suite nulle).

L'exemple suivant montre que le spectre d'un opérateur compact non nul, peut être réduit à $\{0\}$.

Exemple 4.2.11. - Soit $E = L^2(0,1)$, l'espace des fonctions de carrés inté-grables sur $(0,1)$ pour la mesure de Lebesgue, et soit V l'opérateur de Volterra défini par

$$Vf(x) = \int_0^x f(t)dt$$

C'est un opérateur de Hilbert-Schmidt, donc compact. L'équation

$$Vf = \lambda f, \quad \lambda \in \mathbb{C} \setminus \{0\}$$

implique que f est continue, dérivable sur $(0,1)$, et vérifie

$$\lambda f' = f \quad \text{et} \quad f(0) = 0$$

La fonction f est donc identiquement nulle.

EXERCICES

1. Soit A un opérateur compact dans un espace de Hilbert E et soit p un polynôme en une seule variable et sans terme constant. On sait que $p(A)$ est compact. Montrer que si $\lambda \neq 0$ est une valeur propre de A, et n'est pas racine de p alors $p(\lambda)$ est une valeur propre de $p(A)$.

   Solution : Soit $\lambda \neq 0$ une valeur propre de A et soit $v \neq 0$ un vecteur propre associé à λ. On a $Av\lambda v$ et pour tout entier k $A^k = \lambda^k v$, on en déduit que, pour tout polynôme p en une variable sans terme constant, on a $p(A)v = p(\lambda)v$. □
2. Montrer que l'alternative de Fredholm a lieu pour un opérateur A de $\mathcal{L}(E)$, dont une puissance quelconque est un opérateur compact.
3. Soit $I = [a, b]$ un intervalle fermé borné de $\mathbb{R}$ et k une fonction continue sur $I \times I$. Montrer que le spectre de l'opérateur intégral de Fredholm

$$f \mapsto \int_a^x k(x,t)f(t)\,dt, \quad f \in L^2((a,b),dx)$$

   est réduit à $\{0\}$.
4. Pour quelles valeurs de $\mu \in \mathbb{R}$, l'équation

$$f(x) - \mu \int_a^b e^{(x-y)} f(y)\,dy = 1$$

   admet-elle une solution dans $L^2((a,b),dx)$?

   Indication : L'opérateur intégrale A de noyau $k(x,y) = e^{(x-y)}$ est compact (en fait de Hilbert-Schmidt). L'équation $f - \mu Af = 0$ n'admet de solution non nulle que pour $\mu = 1/(b-a)$. Donc, l'équation avec second membre $f - \mu Af = 1$ admet une solution pour tout μ différent de $1/(b-a)$. □
5. Soit K l'opérateur intégral de noyau la fonction k définie sur $[0,1] \times [0,1]$ par $k(x,y) = x - y$.
   Déterminer le spectre de l'opérateur K.
6. Même exercice que 5., où k est défini sur $[0,\pi] \times [0,\pi]$ par

$$k(x,y) = \sin(x)\sin(2y)$$

4.3 Etude spectrale d'un opérateur compact auto-adjoint

Dans ce paragraphe, on va voir que lorsque A est en plus auto-adjoint (non nul), son spectre ne peut être réduit à 0 et qu'un tel opérateur est

diagonalisable, comme c'est le cas pour une matrice hermitienne.

Soient E un espace de Hilbert et A un opérateur auto-adjoint dans $\mathcal{L}(E)$. On sait que

$$\|A\| = \sup_{\|x\|=1} |\langle Ax, x\rangle|$$

Un intérêt des opérateurs compacts est que ce maximum est atteint ; plus précisément, on a

Théorème 4.3.1. *Si A est un opérateur compact auto-adjoint, alors il admet une valeur propre λ telle que $|\lambda| = \|A\|$.*

Démonstration. Posons $a = \|A\|$. On a

$$a = \sup_{\|x\|=1} \langle Ax, x\rangle, \quad \text{ou bien} \quad -a = \inf_{\|x\|=1} \langle Ax, x\rangle$$

En considérant, si nécessaire, $-A$ à la place de A, on peut toujours supposer que

$$a = \sup_{\|x\|=1} \langle Ax, x\rangle$$

Il existe alors, une suite (x_n) d'éléments de E, telle que

$$\forall n \in \mathbb{N},\ \|x_n\| = 1 \quad \text{et} \quad \lim_{n\to\infty} \langle Ax_n, x_n\rangle = a$$

L'opérateur A étant compact, on peut extraire une sous-suite (x_{n_k}) dont l'image par A est convergente. Posons

$$y = \lim_{k\to\infty} Ax_{n_k}$$

On a, bien sûr, $\|y\| \leq a$. Montrons que $Ay = ay$. En effet,

$$\|Ax_{n_k} - ax_{n_k}\|^2 = \|Ax_{n_k}\|^2 + a^2 - 2a\Re e\langle Ax_{n_k}, x_{n_k}\rangle$$

Comme A est auto-adjoint $\Re e\langle Ax_{n_k}, x_{n_k}\rangle = \langle Ax_{n_k}, x_{n_k}\rangle$; en faisant tendre k vers l'infini dans ce qui précède, il vient

$$\lim_{k\to\infty} \|Ax_{n_k} - ax_{n_k}\|^2 = \|y\|^2 + a^2 - 2a^2 = \|y\|^2 - a^2 \leq 0$$

On a donc, en fait, l'égalité

$$\lim_{k\to\infty} \|Ax_{n_k} - ax_{n_k}\|^2 = 0$$

L'opérateur A étant continu, on en déduit que

$$\lim_{k\to\infty} A(Ax_{n_k} - ax_{n_k}) = 0$$

c'est-à-dire que $Ay - ay = 0$, ce qui est le résultat désiré. □

Remarque 4.3.2. - On sait que pour un opérateur auto-adjoint A, le spectre est inclus dans l'intervalle $[-\|A\|, \|A\|]$. Le théorème précédent précise que si A est en plus compact, alors l'une au moins des extrémités de cet intervalle est valeur propre; celle-ci est évidemment la plus grande en valeur absolue.

Soient A un opérateur compact auto-adjoint et (λ_n), $n \geq 1$, la suite de ses valeurs propres, on sait (théorème 2.9) que c'est ou bien une suite finie ou bien elle tend vers 0. On suppose, dans toute la suite, que les valeurs propres sont ordonnées de façon que $(|\lambda_n|)$ soit une suite décroissante. Pour tout $n \geq 1$, on désigne par E_n le sous-espace propre correspondant à la valeur propre λ_n et par P_n la projection orthogonale sur E_n.

Proposition 4.3.3. *Soit $F_n = E_1 \oplus \ldots \oplus E_n$. Alors pour tout n*

$$|\lambda_{n+1}| = \sup_{\substack{x \in F_n^\perp \\ \|x\|=1}} |\langle Ax, x\rangle|$$

Démonstration. L'opérateur A laisse stable F_n, il laisse donc stable son orthogonal $F_n^\perp$ et induit un opérateur $R_n \in \mathcal{L}(F_n^\perp)$

$$\begin{aligned} R_n x &= Ax, \quad \forall x \in F_n^\perp \\ R_n x &= 0, \quad\ \ \forall x \in F_n \end{aligned}$$

L'opérateur R_n est compact et auto-adjoint et ses valeurs propres sont les (λ_j), $j \geq n+1$. On en déduit, grâce au théorème 3.1, que sa valeur propre, la plus grande en valeur absolue, et qui n'est autre que λ_{n+1} vérifie

$$|\lambda_{n+1}| = \|R_n\| = \sup_{\substack{x \in F_n^\perp \\ \|x\|=1}} |\langle R_n x, x\rangle|$$

ce qui fournit le résultat voulu. □

Les sous-espaces propres de A étant deux à deux orthogonaux, on peut considérer leur somme directe hilbertienne que nous noterons F (proposition 4.13, chapitre 1). C'est l'espace de Hilbert des éléments x de la forme

$$x = \sum_{n=1}^{\infty} x_n, \ x_n \in E_n, \quad \text{avec} \quad \|x\| = \left(\sum_{n=1}^{\infty} \|x_n\|^2\right)^{\frac{1}{2}} < \infty$$

et comme A est continu, on a aussi pour tout x dans F

$$Ax = \sum_{n=1}^{\infty} \lambda_n x_n, \quad \text{et} \quad \|Ax\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2 \|x_n\|^2$$

Nous allons voir que ces dernières relations restent vraies pour tout élément x de E.

Théorème 4.3.4. *Tout élément x de E s'écrit de façon unique sous la forme*

$$x = \sum_{n=1}^{\infty} x_n + x_0$$

où, pour tout entier $n \geq 1$, x_n désigne la projection orthogonale de x sur E_n et x_0 sa projection orthogonale sur $\ker(A)$. *De plus, on a*

$$Ax = \sum_{n=1}^{\infty} \lambda_n x_n, \quad \forall x \in E$$

Remarque 4.3.5. - Ce théorème se traduit par les relations suivantes

$$E = F \oplus \ker(A) \quad \text{et} \quad A = \sum_{n=1}^{\infty} \lambda_n P_n$$

où F est la somme directe hilbertienne des sous-espaces propres (E_n) et où P_n est l'opérateur de projection orthogonale sur E_n.

Démonstration. Soit $F_n = E_1 \oplus \ldots \oplus E_n$ et soit A_n et R_n les opérateurs définis par

$$A_n = \sum_{i=1}^{n} \lambda_i P_i, \qquad R_n = A - A_n$$

R_n est l'opérateur induit par A sur $F_n^{\perp}$, que nous avons introduit dans la preuve de la proposition précédente. On a alors

$$\lim_{n \to \infty} \|A - A_n\| = \lim_{n \to \infty} \|R_n\| = \lim_{n \to \infty} |\lambda_{n+1}| = 0$$

ce qui traduit le fait que la série $\sum_{n \geq 1} \lambda_n P_n$ est convergente et a pour somme A. On en déduit que pour tout $x \in E$,

$$Ax = \sum_{n=1}^{\infty} \lambda_n x_n \quad \text{où} \quad x_n = P_n x$$

D'autre part, $\ker(A)$ étant orthogonal à E_n, pour tout $n \geq 1$, il est orthogonal à F et on a l'inclusion $\ker(A) \subset F^{\perp}$; montrons qu'il y a égalité. Soit $x \in F^{\perp}$, pour tout n, $A_n x = 0$, en passant à la limite sur n, on en déduit que $Ax = 0$, c'est-à-dire que x appartient à $\ker(A)$, et on a donc $F^{\perp} = \ker(A)$. L'égalité $E = F \oplus \ker(A)$ en est une conséquence. □

Corollaire 4.3.6. *Si E est un espace séparable, il existe une base hilbertienne de E formée de vecteurs propres de A.*

Démonstration. Si E est séparable, le noyau de A l'est aussi. Comme les espaces propres E_n, $n \geq 1$, sont de dimension finie, il suffit, alors, de choisir une base orthonormée dans chacun des sous-espaces E_n et une base orthonormée dans $\ker(A)$. □

ALTERNATIVE DE FREDHOLM

Soit A un opérateur compact et auto-adjoint dans un espace de Hilbert E. Soit $\lambda \neq 0$, un nombre complexe et considérons l'équation

$$Ax - \lambda x = y \qquad (*)$$

où y est donné dans E et x est l'inconnue. Deux cas se présentent, selon que λ est dans le spectre de A ou non.

PREMIER CAS : $\lambda \notin \sigma_p(A)$

Dans ce cas l'équation $(*)$ admet une unique solution $x \in E$, car $A - \lambda I$ est inversible. Grâce au théorème 3.4, on peut exprimer la solution à l'aide d'un développement en série suivant les vecteurs propres de A.

Soit $\{\lambda_n;\ n \geq 1\}$ la suite des valeurs propres non nulles de A et (E_n) les sous-espaces propres associés. Le théorème 3.4 permet d'écrire

$$y = \sum_1^\infty y_n + y_0, \quad \text{où} \quad y_n \in E_n \quad \text{et} \quad y_0 \in \ker A$$

$$x = \sum_1^\infty x_n + x_0, \quad \text{où} \quad x_n \in E_n \quad \text{et} \quad x_0 \in \ker A$$

En remplaçant x et y par ces expressions dans l'équation $(*)$ et en identifiant, on obtient $x_n = y_n/(\lambda_n - \lambda)$ et $x_0 = -y_0/\lambda$; la solution x est donc donnée par

$$x = \sum_{n=1}^\infty \frac{y_n}{\lambda_n - \lambda} - \frac{y_0}{\lambda}$$

ce qui s'écrit

$$x = -\frac{y}{\lambda} + \frac{1}{\lambda}\sum_{n=1}^\infty \frac{\lambda_n}{\lambda_n - \lambda} y_n$$

DEUXIÈME CAS : $\lambda \in \sigma_p(A)$

Soit $\lambda = \lambda_k$. L'opérateur $A - \lambda_k I$, restreint au sous-espace de Hilbert $E_k^\perp$ est injectif ; il est donc surjectif car A est compact, et on a

$$(A - \lambda_k I)(E_k^\perp) = E_k^\perp$$

On en déduit que l'équation $(*)$ admet une solution si, et seulement si, y est orthogonal au sous-espace propre, E_k, associé à la valeur propre λ_k. Dans ce cas, l'équation admet une infinité de solutions, deux d'entr'elles diffèrent par un élément quelconque de E_k.

On obtient le développement en série des solutions en procédant comme dans le premier cas. Si

$$y = \sum_{n \neq k}^{\infty} y_n + y_0, \quad x_n \in E_n \quad \text{et} \quad y_0 \in \ker A$$

on vérifie que la solution générale x s'écrit

$$x = \sum_{n \neq k}^{\infty} \frac{y_n}{\lambda_n - \lambda_k} - \frac{y_0}{\lambda_k} + x_k,$$

ou encore

$$x = -\frac{y}{\lambda_k} + \frac{1}{\lambda_k} \sum_{n \neq k} \frac{\lambda_n}{\lambda_n - \lambda_k} y_n + x_k$$

où x_k est un élément arbitraire de $\ker(A - \lambda_k I)$.

Le cas : $\lambda = 0$

L'égalité $\overline{(\Im m A)} = \ker A^{\perp}$ montre que s'il existe $x \in E$ tel que $Ax = y$, alors nécessairement y est orthogonal à $\ker A$, mais cette condition n'est pas suffisante pour assurer l'existence d'une telle solution x. On doit supposer de plus que

$$\sum_{n=1}^{\infty} \frac{\|y_n\|^2}{\lambda_n^2} < \infty$$

Dans ce cas il existe une infinité de solutions, dont la forme générale est donnée par

$$x = \sum_{n=1}^{\infty} \frac{y_n}{\lambda_n} + x_0$$

où x_0 est un élément arbitraire de $\ker A$.

Exemple 4.3.7. - On considère, sur l'espace de Hilbert $L^2[0, \pi]$, l'opérateur intégral K dont le noyau est

$$k(x, y) = \cos(x + y)$$

C'est un opérateur de rang 2 auto-adjoint. On vérifie rapidement que ses valeurs propres non nulles sont

$$\lambda_1 = \frac{\pi}{2} \quad \text{et} \quad \lambda_2 = -\frac{\pi}{2}$$

Elles sont simples et les fonctions propres associées sont

$$\phi_1(x) = \sqrt{\frac{2}{\pi}}\cos x \quad \text{et} \quad \phi_2(x) = \sqrt{\frac{2}{\pi}}\sin x$$

Le noyau $\ker(K)$ est formé des éléments de $L^2[0,\pi]$ orthogonaux à ϕ_1 et à ϕ_2.

On se propose de résoudre, dans $L^2[0,\pi]$, l'équation

$$Kf - \lambda f = -x$$

• Si $\lambda = 0$, cette équation n'a pas de solution, car son second membre n'est pas dans l'image de l'opérateur K.
• Si $\lambda \neq 0$ et si λ n'est pas l'une des deux valeurs propres λ_1 et λ_2, l'équation admet une unique solution donnée par

$$f(x) = \frac{4}{\lambda(\pi - 2\lambda)}\cos x - \frac{2\pi}{\lambda(2\lambda+\pi)}\sin x + \frac{x}{\lambda}$$

• Enfin, si $\lambda = \lambda_1$ ou $\lambda = \lambda_2$, l'équation n'a pas de solution car son second membre n'est orthogonal ni à ϕ_1, ni à ϕ_2.

EXERCICES

1. Les notations étant celles de la proposition 3.3, montrer que les opéra-teurs P_n et A commutent et que $R_n = A - P_nA$.

2. Soient E_1 et E_2 deux espaces de Hilbert et A un opérateur compact de E_1 dans E_2. L'adjoint de A est l'opérateur compact A^* de E_2 dans E_1 défini par

$$\langle Ax, y\rangle = \langle x, A^*y\rangle, \quad \forall\, x \in E_1,\ y \in E_2$$

(a) Vérifier que A^* est bien un opérateur compact de E_2 dans E_1.
(b) Montrer que A^*A est compact de E_1 dans lui-même et positif.
On désigne par λ_n^2 ses valeurs propres, chacune répétée un nombre de fois égal à sa multiplicité et on suppose qu'elles sont ordonnées de façon que (λ_n) soit décroissante. Soit (e_n) les fonctions propres associées et soit $f_n = (1/\lambda_n)Ae_n$.
(c) Montrer que (f_n) est une suite orthogonale dans E_2 et que, pour tout x dans E_1,

$$Ax = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle f_n$$

3 On considère l'espace $L^2(0, 2\pi)$, muni du produit scalaire

$$\langle f, g\rangle = \int_0^{2\pi} f(x)\overline{g(x)}\,dx$$

Soit A l'opérateur intégral défini pour $f \in L^2(0, 2\pi)$ par

$$Af(x) = x \int_0^{2\pi} f(y) \cos y \, dy + \cos x \int_0^{2\pi} y f(y) \, dy$$

On pose $u_1(x) = x$ et $u_2(x) = \cos x$.
(1) Montrer que A est un opérateur auto-adjoint de fini.
(2) Montrer que u_1 et u_2 sont deux éléments orthogonaux et en déduire une description de l'image de A et de son noyau.
(3) Déterminer les valeurs propres de A et les sous-espaces propres correspondants.
(4) On pose $v(x) = \sin x$. Discuter, suivant la valeur de $\lambda \in \mathbb{C}$, le nombre de solutions $u \in L^2(0, 2\pi)$ de l'équation

$$\lambda u - Au = v$$

Solution : Une intégration par parties permet de voir que $\langle u_1, u_2 \rangle = 0$. D'autre part, l'expression même de Af montre que A est un opérateur intégral dont le noyau est la fonction définie par $k(x, y) = x \cos y + y \cos x$. Cette fonction étant symétrique réelle, l'opérateur A est auto-adjoint. Il est clair que l'image de A est engendrée par u_1 et u_2, ces deux éléments forment donc une base orthogonale de l'image de A et par suite $\ker(A) = \Im m(A)^\perp$ (voir le théorème 4.6 du chapitre III). On en déduit que A est de rang 2. Pour déterminer les valeurs propres de A, on remarque que si ϕ est fonction propre associée à une valeur propre λ, alors ϕ appartient à l'image de A et est donc forcément de la forme $\phi = au_1 + bu_2$. On montre facilement que

$$Au_1 = \frac{(2\pi)^3}{3} u_2 \quad \text{et} \quad Au_2 = \pi u_1$$

L'égalité $A\phi = \lambda\phi$ montre alors que

$$\lambda \in \left\{ -4\pi^2/\sqrt{6}, \ 0, \ 4\pi^2/\sqrt{6} \right\}$$

De plus, pour $\lambda_1 =, -4\pi^2/\sqrt{6}$, on trouve que ϕ est de la forme $\phi = a(u_1 + \lambda_1 u_2)$, où $a \in \mathbb{C}$. Il en résulte que le sous-espace propre associé à la valeur propre λ_1 est de dimension 1 et engendré par $\phi_1 = u_1 + \lambda_1 u_2$. De même le sous-espace propre associé à la valeur propre $\lambda_2 = 4\pi^2/\sqrt{6}$ est de dimension 1 et engendré par $\phi_2 = u_1 + \lambda_2 u_2$. Quant au sous-espace propre associé à la valeur propre 0, c'est $\ker(A)$ qui est de dimension infinie. Il reste à discuter, selon λ, le nombre de solutions de l'équation intégrale $\lambda u - Au = v$: On vérifie que $\langle v, u_1 \rangle = -2\pi$ et $\langle v, u_2 \rangle = 0$ et par suite, si $\lambda = \lambda_1$ ou si $\lambda = \lambda_2$, l'équation n'admet pas de solution. On vérifie aussi que v n'appartient pas à l'image de A et par suite, pour $\lambda = 0$ l'équation n'admet pas de solution. Enfin, pour toute autre valeur de λ l'équation admet une solution et une seule. □

Chapitre 5

Problème de Sturm-Liouville

Plusieurs équations de la physique mathématique telles que l'équation des ondes, l'équation de Laplace[1], l'équation de la chaleur, l'équation de Schrödinger[2] etc..., peuvent être traitées grâce à la méthode de séparation des variables qui ramène ces équations aux dérivées partielles, à des équations différentielles linéaires du second ordre de la forme

$$\alpha(x)u'' + \beta(x)u' + \gamma(x)u = \lambda u, \quad \text{où} \quad \lambda \in \mathbb{C}$$

équations dont on cherche les solutions u satisfaisant à des conditions imposées par le problème physique étudié. Dans beaucoup de cas, la méthode de la variation des constantes de Lagrange ramène la résolution de l'équation précédente à celle d'une équation intégrale de Fredholm à noyau hermitien, pour laquelle on peut appliquer les développements du chapitre IV.

5.1 Opérateur à Noyau hermitien continu

Dans ce paragraphe, on considère le cas particulier où A est un opérateur à noyau. Soit I un intervalle de $\mathbb{R}$ et k une fonction de $I \times I$ à valeurs complexes, mesurable et de carré intégrable pour la mesure de Lebesgue. On suppose que k est hermitienne, c'est-à-dire vérifiant $k(x,y) = \overline{k(y,x)}$. On désigne par K l'opérateur intégral de noyau k défini, pour f dans $L^2(I)$,

[1]Pierre Simon de LAPLACE, (1749-1827), se distingua par de nombreux travaux d'astronomie, de mathématiques et de physique. Il appliqua l'analyse mathématique à la mécanique céleste et à la théorie des probabilités. C'est dans son ouvrage "Théorie analytique des probabilités" qu'il introduisit, en 1812, la transformation qui porte son nom, pour caractériser diverses lois de probabilité.

[2] Erwin Schrödinger (1887-1961) est un physicien autrichien. Il a donné, en 1926, une formalisation nouvelle de la théorie quantique, introduisant en particulier l'équation fondamentale (qui porte son nom), à la base de tous les calculs de la spectroscopie. Il a reçu le prix Nobel en 1933.

par

$$Kf(x) = \int_I k(x,y)f(y)\,dy$$

On sait que K est un opérateur de Hilbert-Schmidt (donc compact) et auto-adjoint.

Soit (λ_n) la suite des valeurs propres non nulles de K, chacune répétée autant de fois que sa multiplicité, et rangées de façon que $(|\lambda_n|)$ soit une suite décroissante, et soit (ϕ_n) la suite des fonctions propres associées, qu'on suppose normalisées. La proposition 1.22, chapitre IV, dit que

$$k = \sum_{n=1}^{\infty} \lambda_n \phi_n \otimes \overline{\phi_n} \qquad \text{égalité dans } L^2(I \times I)$$

et la norme de Hilbert-Schmidt de l'opérateur K est donnée par

$$|||\, K \,|||^2 = \|k\|^2_{L^2(I\times I)} = \sum_{n=1}^{\infty} \lambda_n^2$$

L'équation intégrale de Fredholm s'écrit ici

$$\int_I k(x,y)u(y)\,dy - \lambda u = g$$

où g est donnée et u la fonction à chercher. Cette équation est appelée équation de Fredholm de *première espèce* si $\lambda = 0$, et équation de Fredholm de *seconde espèce* si $\lambda \neq 0$.

• On suppose, dans toute la suite, que I est un intervalle fermé borné $[a,b]$ et que k **est continu sur** $[a,b] \times [a,b]$. Nous allons voir que, dans ces conditions, les séries qui figurent dans le paragraphe 3.6 du chapitre IV convergent, non seulement au sens de la norme de $L^2[a,b]$, mais aussi uniformément et absolument, dès que le second membre g est une fonction continue sur l'intervalle $[a,b]$.

Théorème 5.1.1. *L'opérateur K est une application compacte de l'espace de Hilbert $L^2[a,b]$ dans l'espace $(C[a,b], \|\ \|_\infty)$.*

Démonstration. Soit f dans $L^2[a,b]$. Pour x_1 et x_2 dans $[a,b]$, l'inégalité de Cauchy-Schwarz donne

$$|Kf(x_1) - Kf(x_2)|^2 \leq \int_a^b |k(x_1,y) - k(x_2,y)|^2\,dy\, \|f\|^2$$

Il en résulte que si (f_n) est une suite bornée d'éléments de $L^2[a,b]$ ($\|f_n\| \leq M$, pour tout n), la suite (Kf_n) est uniformément bornée et équicontinue sur $[a,b]$. D'après le théorème d'Ascoli (voir l'annexe), il est possible d'en extraire une sous-suite qui converge uniformément sur l'intervalle $[a,b]$. □

Remarque 5.1.2. - Ce qui précède montre en outre que, pour toute f dans $L^2[a,b]$, Kf est une fonction continue sur $[a,b]$. En particulier, puisque $K\phi_n = \lambda_n\phi_n$, les fonctions propres associées aux valeurs propres non nulles de l'opérateur K sont des fonctions continues sur l'intervalle $[a,b]$.

Théorème 5.1.3. *Pour toute fonction f de $L^2[a,b]$, on a*

$$Kf(x) = \sum_{n=1}^{\infty} \lambda_n \langle f, \phi_n\rangle \phi_n(x) \tag{5.1}$$

où la série converge absolument et uniformément sur $[a,b]$.

Démonstration. La relation $\lambda_n\phi_n(x) = K\phi_n(x)$ peut s'écrire $\overline{\lambda}_n\overline{\phi_n(x)} = \langle k(.,x), \phi_n\rangle$. D'après l'inégalité de Bessel, il vient

$$\sum_{n=1}^{\infty} \lambda_n^2 |\phi_n(x)|^2 \leq \int_a^b |k(x,y)|^2\, dy$$

et par suite, l'inégalité de Cauchy-Schwarz permet d'écrire

$$\left(\sum_{n=p}^{q} |\lambda_n \langle f, \phi_n\rangle \phi_n(x)|\right)^2 \leq \sum_{n=p}^{q} |\langle f, \phi_n\rangle|^2 \int_a^b |k(x,y)|^2\, dy$$

Comme la série $\sum |\langle f, \phi_n\rangle|^2$ est convergente et que k est continue sur $[a,b] \times [a,b]$, de l'inégalité précédente on déduit que la série $\sum \lambda_n \langle f, \phi_n\rangle \phi_n$ est absolument et uniformément convergente sur l'intervalle $[a,b]$. Or cette série converge aussi dans $L^2[a,b]$ vers Kf, celle-ci est donc sa limite uniforme, ce qui termine la preuve du théorème. □

Considérons maintenant l'équation intégrale

$$\lambda u(x) - \int_a^b k(x,y)u(y)\, dy = f(x) \tag{*}$$

où λ est un nombre complexe non nul, f est une fonction continue donnée sur $[a,b]$ et où u est une fonction continue à déterminer.

Théorème 5.1.4. (i) *Si λ n'est pas valeur propre de K, l'équation $(*)$ admet une solution unique donnée par*

$$u(x) = \frac{1}{\lambda} f(x) + \frac{1}{\lambda} \sum_{n=1}^{\infty} \frac{\lambda_n}{\lambda - \lambda_n} \langle f, \phi_n\rangle \phi_n(x) \tag{5.2}$$

où la série converge absolument et uniformément sur $[a,b]$.

(ii) *Si λ est une valeur propre de K, l'équation $(*)$ n'admet de solution que si f est orthogonale au sous-espace propre E_λ correspondant à λ, les solutions sont alors données par*

$$u(x) = \frac{1}{\lambda} f(x) + \frac{1}{\lambda} \sum_{\lambda_n \neq \lambda} \frac{\lambda_n}{\lambda - \lambda_n} \langle f, \phi_n \rangle \phi_n(x) + u_\lambda \tag{5.3}$$

où u_λ est une fonction arbitraire dans le sous-espace propre E_λ ; la convergence de la série du second membre étant absolue et uniforme sur $[a, b]$.

Démonstration. (i) Supposons que λ ne soit pas valeur propre de K. Si l'équation intégrale $(*)$ admet une solution u, celle-ci vérifie

$$\lambda u(x) - f(x) = \int_a^b k(x, y) u(y)\, dy$$

donc d'après le théorème 1.3

$$\lambda u(x) - f(x) = \sum_{n=1}^{\infty} \lambda_n \langle u, \phi_n \rangle \phi_n(x)$$

où la convergence est absolue et uniforme sur $[a, b]$. En multipliant les deux membres par $\overline{\phi_j}(x)$ et en intégrant terme à terme sur l'intervalle $[a, b]$, on obtient l'égalité $\lambda \langle u, \phi_j \rangle - \langle f, \phi_j \rangle = \lambda_j \langle u, \phi_j \rangle$, c'est-à-dire $\langle u, \phi_j \rangle = \langle f, \phi_j \rangle / (\lambda - \lambda_j)$, pour tout j. L'égalité (2) s'en déduit.

Réciproquement, pour toute fonction continue f, la formule (2) ci-dessus fournit une solution de l'équation intégrale $(*)$. Pour le voir, montrons d'abord que la série converge uniformément sur $[a, b]$; on a déjà vu, au cours de la preuve du théorème 1.3, que

$$\sum_{n=1}^{\infty} \lambda_n^2 |\phi_n(x)|^2 \leq \int_a^b |k(x, y)|^2\, dy$$

En utilisant l'inégalité de Cauchy-Schwarz, il vient

$$\begin{aligned} \left| \sum_{n+1}^{\infty} \frac{\lambda_j}{\lambda - \lambda_j} \langle f, \phi_j \rangle \right|^2 &\leq C \sum_{n+1}^{\infty} |\langle f, \phi_j \rangle|^2 \sum_{n+1}^{\infty} \lambda_j^2 |\phi_j(x)|^2 \\ &\leq C \sum_{n+1}^{\infty} |\langle f, \phi_n(x) \rangle|^2 \int_a^b |k(x, y)|^2\, dy \end{aligned}$$

avec $C = \sup_j |\lambda - \lambda_j|^{-1}$. Comme k est continue, le second membre tend vers 0 lorsque n tend vers l'infini et ce indépendamment de x, d'où la convergence uniforme. Maintenant il est facile de vérifier, en calculant $\lambda u - K u$,

que la fonction u, définie par la formule ci-dessus est solution de l'équation intégrale $(*)$.

(ii) Supposons que λ soit valeur propre de K. Soit u une solution de $(*)$ et soit ϕ une fonction propre de K, correspondant à la valeur propre λ, on a

$$0 = \lambda\langle u, \phi\rangle - \langle Ku, \phi\rangle = \langle f, \phi\rangle$$

Ainsi, pour qu'une solution de $(*)$ existe il est nécessaire que f soit orthogonale au sous-espace propre correspondant à la valeur propre λ. D'autre part, on vérifie comme précédemment que la fonction u donnée par (2) est solution et que la convergence de la série est absolue et uniforme sur l'intervalle $[a, b]$. □

Corollaire 5.1.5. *L'unique solution u de l'équation $(*)$ s'écrit*

$$u(x) = -\frac{1}{\lambda}f(x) + \int_a^b R(x, y; \lambda)f(y)\, dy$$

avec

$$R(x, y; \lambda) = -\frac{1}{\lambda^2}k(x, y) + \frac{1}{\lambda^2}\sum_{n=1}^{\infty}\frac{\lambda_n^2}{(\lambda_n - \lambda)}\phi_n(x)\overline{\phi_n}(y)$$

et où la série du second membre est absolument et uniformément convergente sur $[a, b]$.

Démonstration. On a l'égalité

$$\lambda_n/\lambda(\lambda_n - \lambda) = -\lambda_n/\lambda^2 + \lambda_n^2/\lambda^2(\lambda_n - \lambda)$$

et le théorème précédent assure que l'unique solution f de $(*)$ est donnée par

$$f(x) = -\frac{g}{\lambda} - \frac{1}{\lambda^2}\sum_{n=1}^{\infty}\lambda_n\langle g, \phi_n\rangle\phi_n(x) + \sum_{n=1}^{\infty}\frac{\lambda_n^2}{\lambda^2(\lambda_n - \lambda)}\langle g, \phi_n\rangle\phi_n(x)$$

Comme la série

$$\sum_{n=1}^{\infty}\lambda_n\langle g, \phi_n\rangle\phi_n(x)$$

converge absolument et uniformément sur $[a, b]$ vers Kg, on en déduit que

$$f(x) = -\frac{g}{\lambda} - \frac{1}{\lambda^2}\int_a^b k(x, y)g(y)\, dy$$
$$+ \sum_{n=1}^{\infty}\frac{\lambda_n^2}{\lambda^2(\lambda_n - \lambda)}\left(\int_a^b g(y)\overline{\phi_n}(y)\, dy\right)\phi_n(x)$$

Il suffit donc de montrer que la série $\sum_n \lambda_n^2 |\phi_n(y)|^2$ est uniformément convergente sur $[a,b]$, auquel cas la série $\sum_n \lambda_n^2 \phi_n(x)\overline{\phi_n}(y)$ sera aussi absolument et uniformément convergente sur $[a,b] \times [a,b]$, et l'interversion des signes $\sum$ et $\int$, dans le dernier terme du second membre terminera la preuve. A cet effet, considérons la fonction

$$h(x,y) = \int_a^b k(x,t)k(t,y)dt$$

C'est une fonction continue sur $I \times I$ et pour chaque y fixé, on peut appliquer le théorème 1.3 pour affirmer que

$$h(x,y) = \sum_{n=1}^{\infty} \lambda_n^2 \phi_n(x)\overline{\phi_n}(y)$$

où la convergence est absolue et uniforme en $x \in [a,b]$. En particulier la série converge simplement, en tout point (x,y) de $[a,b] \times [a,b]$. Il vient notamment

$$h(x,x) = \sum_{n=1}^{\infty} \lambda_n^2 |\phi_n(x)|^2, \quad \forall\, x \in [a,b]$$

On a là une série à termes positifs de fonctions continues qui converge en tout point de $[a,b]$ vers une fonction continue, le théorème de Dini[3] (voir l'annexe) assure que la convergence de la série est en fait uniforme sur l'intervalle $[a,b]$. □

Notons que, en général, la série $\sum \lambda_n \phi_n(x)\overline{\phi_n}(y)$ ne converge pas. Nous allons donner, dans ce qui suit, une condition sur le noyau k qui assure sa convergence uniforme sur l'intervalle produit.

Définition 5.1.6. *On dit qu'un noyau k, continu sur $[a,b] \times [a,b]$, est de type positif s'il vérifie pour tout f dans $L^2[a,b]$*

$$\int_a^b \int_a^b k(x,y) f(y)\overline{f(x)}\; dydx \geq 0,$$

Compte tenu de la définition 4.16, chapitre III, k est de type positif si, et seulement si, l'opérateur K est positif.

Proposition 5.1.7. *Tout noyau k de type positif vérifie, pour tout (x,y) dans $[a,b] \times [a,b]$,*

(a) $k(x,x) \geq 0$

[3]Ulisse DINI (1845-1948), est un mathématicien italien, dont la statue, près de la Scuola Normale di Pisa, est régulièrement décorée par les étudiants pisans pour qui Pise n'est pas réduite à sa tour.

(b) $k(x,y) = \overline{k(y,x)}$

Démonstration. Si k est de type positif, l'opérateur intégral K, de noyau k, est positif donc auto-adjoint et son noyau k est hermitien, c'est-à-dire satisfait à la propriété (b). Supposons maintenant qu'il existe x_0 dans l'intervalle $[a,b]$, tel que $k(x_0,x_0) < 0$. Il existe c et d, avec $a \leq c < x_0 < d \leq b$, tels que

$$\Re e(k(x,y)) \leq 0, \quad \text{pour} \quad (x,y) \in [c,d] \times [c,d]$$

En prenant pour f la fonction caractéristique de l'intervalle $[c,d]$, on obtient

$$0 \leq \int_a^b \int_a^b k(x,y) f(y) \overline{f(x)}\, dydx = \int_c^d \int_c^d k(x,y)\, dydx < 0$$

ce qui est absurde. Le noyau k satisfait donc la propriété (a). □

Notons que si le noyau k est de type positif, l'opérateur intégral K qui lui est associé est positif et ses valeurs propres sont positives ou nulles. Comme précédemment nous désignerons par (λ_n) et (ϕ_n) les valeurs propres et fonctions propres de K.

Théorème 5.1.8. (de Mercer) *Si k est continu sur $[a,b] \times [a,b]$ et de type positif, alors pour tout (x,y) dans $[a,b] \times [a,b]$*

$$k(x,y) = \sum_{n=1}^{\infty} \lambda_n \phi_n(x) \overline{\phi_n}(y)$$

où la série converge absolument et uniformément sur $[a,b] \times [a,b]$.

Démonstration. L'opérateur K est positif et donc toutes ses valeurs propres le sont. Pour n un entier posons

$$R_n(x,y) = k(x,y) - \sum_{j=1}^{n} \lambda_j \phi_j(x) \overline{\phi_j}(y)$$

Le noyau R_n est également continu et de type positif car, pour toute fonction f dans $L^2[a,b]$,

$$\langle R_n f, f \rangle = \sum_{n+1}^{\infty} \lambda_j |\langle f, \phi_j \rangle|^2 \geq 0$$

Grâce à la proposition 1.7, on en déduit que $R_n(x,x) \geq 0$. En revenant à la définition de R_n, cette inégalité implique que, pour tout n,

$$\sum_{j=1}^{n} \lambda_j |\phi_j(x)|^2 \leq k(x,x) \quad \text{et par suite} \quad \sum_{j=1}^{\infty} \lambda_j |\phi_j(x)|^2 \leq k(x,x)$$

Cela montre que la série $\sum \lambda_j|\phi_j(x)|^2$ est convergente pour tout x dans $[a,b]$. Utilisant l'inégalité de Cauchy-Schwarz, il vient

$$\sum_{j=p}^{q} \lambda_j|\phi_j(x)\overline{\phi_j}(y)| \leq \left(\sum_{j=p}^{q} \lambda_j|\phi_j(x)|^2\right)^{\frac{1}{2}} \left(\sum_{j=p}^{q} \lambda_j|\phi_j(y)|^2\right)^{\frac{1}{2}}$$
$$\leq \sqrt{k(x,x)}\left(\sum_{j=p}^{q} \lambda_j|\phi_j(y)|^2\right)^{\frac{1}{2}}$$

Comme la fonction k est bornée sur $[a,b]\times[a,b]$, pour tout y fixé la convergence de la série $\sum \lambda_j\phi_j(x)\overline{\phi_j}(y)$ est absolue et uniforme en x sur $[a,b]$. Or, cette même série converge dans $L^2([a,b]\times[a,b])$ vers $k(x,y)$ (proposition 1. 22, chapitre IV) ; il en résulte que

$$\sum_{j=1}^{\infty} \lambda_j\phi_j(x)\overline{\phi_j}(y) = k(x,y), \quad \text{où } \ x,y\in[a,b]$$

En particulier, $k(x,x) = \sum_{j=1}^{\infty} \lambda_j|\phi_j(x)|^2$. Le théorème de Dini assure que la convergence de cette dernière série est uniforme sur $[a,b]$. D'après l'inégalité de Cauchy-Schwarz,

$$\left(\sum_{j=p}^{q} \lambda_j\phi_j(x)\overline{\phi_j}(y)|\right)^2 \leq \left(\sum_{j=p}^{q} \lambda_j|\phi_j(x)|^2\right)\left(\sum_{j=p}^{q} \lambda_j\overline{\phi_j}(y)|^2\right)$$

si bien que la série $\sum \lambda_j\phi_j(x)\overline{\phi_j}(y)$ est absolument et uniformément convergente sur $[a,b]\times[a,b]$ et sa somme est égale à $k(x,y)$. □

Corollaire 5.1.9. (Formule de la Trace) *Si k est un noyau continu sur $[a,b]\times[a,b]$ et de type positif, alors*

$$\int_a^b k(x,x)\,dx = \sum_{n=1}^{\infty} \lambda_n$$

Le second membre est appelé la trace de l'opérateur A.

Démonstration. La série $\sum \lambda_n|\phi_n(x)|^2$ converge uniformément sur $[a,b]$ vers $k(x,x)$, d'après le théroème de Mercer. On peut donc intégrer terme à terme cette série et le corollaire s'en déduit. □

EXERCICES

1. Soit l'espace $L^2([-\pi,\pi]\times[-\pi,\pi])$ des fonctions mesurables et de carré intégrable relativement à la mesure de Lebesgue. Le produit scalaire et la norme sont donnés par

$$\langle f,g\rangle = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(x,y)\overline{g(x,y)}\,dxdy, \quad \|f\|^2 = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}|f(x,y)|^2\,dxdy$$

Montrer que dans cet espace, la suite $(\Phi_{n,m})$ définie par

$$\Phi_{n,m}(x,y) = \frac{1}{4\pi^2}e^{(inx+imy)}, \quad \text{avec} \quad n,m\in\mathbb{Z}$$

constitue une base hilbertienne et écrire le développement d'une fonction f de $L^2([-\pi,\pi]\times[-\pi,\pi])$ suivant la base $(\Phi_{n,m})$.
Solution : Il suffit de se reporter au lemme 1.23, chapitre IV. □

2. Soit K un opérateur intégral à noyau k continu et hermitien sur $[a,b]$. On pose

$$k_2(x,y) = \int_a^b k(x,t)k(t,y)\,dt$$

(a) Les notations étant celles du paragraphe 1, montrer que

$$k_2(x,y) = \sum_{n=1}^{\infty}\lambda_n^2\phi_n(x)\overline{\phi_n(y)}$$

où l'on précisera la nature de la convergence de la série.
(b) On définit par récurrence k_p par

$$k_1 = k \quad \text{et} \quad k_{p+1}(x,y) = \int_a^b k_p(x,t)k(t,y)\,dt, \quad p\geq 1$$

et on désigne par K_p l'opérateur intégral de noyau k_p. Montrer que les valeurs propres de K_p sont les (λ_n^p) et que les fonctions propres associées sont les (ϕ_n). En déduire que

$$k_p(x,y) = \sum_{n=1}^{\infty}\lambda_n^p\phi_n(x)\overline{\phi_n(y)}$$

(c) Montrer que lorsque $p>1$ la série ci-dessus est absolument et uniformément convergente sur $[a,b]$. En déduire que, pour tout $p>1$,

$$\int_a^b k_p(x,x)\,dx = \sum_{n=1}^{\infty}\lambda_n^p$$

Que peut-on dire lorsque $p=1$?

(d) Pour quelles valeurs de l'entier p, le noyau k_p est (dans tous les cas) de type positif ?

Solution : (a) Il suffit de se reporter à la preuve du corollaire 1.5 de ce chapitre. Il en ressort que la série définissant k_2 converge uniformément sur $[a,b] \times [a,b]$. La même preuve permet de montrer (b) par récurrence. La convergence uniforme permet d'intégrer les séries terme à terme, ce qui démontre (c). Le noyau k_p est de type positif si, et seulement si, les valeurs propres de l'opérateur associé K_p sont positives. Ceci aura toujours lieu si p est un entier pair. □

3. Soit k un noyau hermitien sur $]a,b[\times]a,b[$. On suppose qu'il existe une constante M telle que

$$\sup_x \int_a^b |k(x,y)|^2\, dy < M < \infty$$

et soit K l'opérateur intégral de noyau k. Montrer que les séries (1) et (2) des théorèmes 1.3 et 1.4 convergent absolument et uniformément sur $[a,b]$ (cela généralise les théorèmes 1.3 et 1.4 au cas où k n'est pas forcément continu). Montrer que, sur $]0,1[$, la fonction k suivante satisfait à la propriété ci-dessus

$$k(x,y) = \begin{cases} \operatorname{Log} x, & \text{si } 0 < y \leq x < 1\,; \\ \operatorname{Log} y, & \text{si } 0 < x \leq y < 1. \end{cases}$$

Solution : Les hypothèses montrent que le noyau k est dans l'espace $L^2\big([a,b] \times [a,b]\big)$. L'opérateur K associé est donc un opérateur de Hilbert-Schmidt. En suivant la preuve du théorème 1.3 et avec les mêmes notations, il vient

$$\sum_{n=1}^{\infty} \lambda_n^2 |\phi_n(x)|^2 = \int_a^b |k(x,y)|^2\, dy \leq M^2, \quad \forall\, x \in [a,b]$$

On en déduit que, pour f dans $L^2[a,b]$ et $p < q$,

$$\begin{aligned} \left(\sum_{n=p}^{q} |\lambda_n \langle f, \phi_n\rangle \phi_n(x)|\right)^2 &\leq \sum_{n=p}^{q} |\langle f, \phi_n\rangle|^2 \int_a^b |k(x,y)|^2\, dy \\ &\leq M \sum_{n=p}^{q} |\langle f, \phi_n\rangle|^2 \end{aligned}$$

Le dernier terme est indépendant de la variable x et tend vers 0 lorsque p et q tendent vers l'infini. Cela prouve que la série

$$\sum \lambda_n \langle f, \phi_n\rangle \phi_n$$

converge uniformément sur $[a, b]$ et sa limite est Kf. Ainsi, on a obtenu l'analogue du théorème 1.3. L'analogue du théorème 1.4 s'en déduit naturellement. □

5.2 Opérateur différentiel du second ordre

Soient α_0, α_1 et α_2 des fonctions continues sur un intervalle I à valeurs réelles. On suppose α_0 strictement positive dans I.

On désigne par $\mathcal{L}$ l'opérateur différentiel linéaire et du second ordre qui, à une fonction u dans l'espace $C^2(I)$ des fonctions deux fois continûment dérivables sur I, associe

$$\mathcal{L}u = \alpha_0 u'' + \alpha_1 u' + \alpha_2 u$$

On rappelle le théorème d'existence et d'unicité de Picard qui dit que, pour chaque choix de x_0 dans I, de (ξ_1, ξ_2) dans $\mathbb{C} \times \mathbb{C}$ et de f dans $C(I)$, le problème

$$\begin{cases} \mathcal{L}u = f \\ u(x_0) = \xi_1, \ u'(x_0) = \xi_2 \end{cases}$$

admet une unique solution u dans $C^2(I)$.

Il en résulte que :

• l'ensemble des solutions de l'équation différentielle homogène $\mathcal{L}u = 0$ est un sous-espace vectoriel de $C^2(I)$ de dimension deux.

• Si u_0 est une solution particulière de l'équation $\mathcal{L}u = f$ et si u_1 et u_2 sont deux solutions linéairement indépendantes de l'équation homogène $\mathcal{L}u = 0$, alors, la solution générale de l'équation $\mathcal{L}u = f$ est de la forme

$$u = c_1 u_1 + c_2 u_2 + u_0$$

où c_1 et c_2 sont deux constantes arbitraires.

Exemple 5.2.1. - La solution générale de l'équation $u'' + u = x$ est donnée par

$$u(x) = c_1 \sin x + c_2 \cos x + x$$

Soient u et v dans $C^2(I)$. Leur wronskien est défini par

$$W(u, v) = uv' - vu'$$

Deux solutions u_1 et u_2 de l'équation homogène $\mathcal{L}u = 0$ sont linéairement indépendantes si, et seulement si, leur wronskien est non nul. Notons à ce propos que si $W(u_1, u_2)$ est non nul en un point t_0 de I, alors il ne s'annule

en aucun point de cet intervalle, cela est une conséquence du théorème d'existence et d'unicité. C'est aussi une conséquence de la formule suivante, dite parfois formule d'Abel[4] : Quels que soient x et y dans l'intervalle I

$$W(u_1, u_2)(x) = W(u_1, u_2)(y) \exp\Big(-\int_y^x \frac{\alpha_1(t)}{\alpha_0(t)} dt\Big) \tag{5.4}$$

De cette relation, dont la preuve est présentée en exercice, on déduit que si u_1 est une solution connue de $\mathcal{L}u = 0$, qui ne s'annule pas sur I, alors une deuxième solution de cette équation est donnée par

$$u_2(x) = u_1(x) \int_\tau^x u_1^{-2}(t) \exp\Big(-\int_\tau^t \frac{\alpha_1(s)}{\alpha_0(s)} ds\Big) dt, \quad (\tau \in I) \tag{5.5}$$

Enfin, la connaissance de deux solutions linéairement indépendantes u_1 et u_2 de $\mathcal{L}u = 0$ permet de construire une solution particulière de l'équation avec second membre $\mathcal{L}u = f$ (voir exercice 2).

On suppose dans la suite que α_0 est dans $C^2(I)$ et que α_1 est dans l'espace $C^1(I)$.

Définition 5.2.2. (L'opérateur adjoint) *L'adjoint formel de l'opérateur $\mathcal{L}$, noté $\mathcal{L}^+$, est l'opérateur défini par*

$$\mathcal{L}^+ u = (\alpha_0 u)'' - (\alpha_1 u)' + \alpha_2 u$$

L'adjoint formel de $\mathcal{L}$ est en général différent de $\mathcal{L}$, il lui est rattaché par la relation suivante, dite identité de Lagrange[5]

Proposition 5.2.3. *Quels que soient u et v dans $C^2(I)$*

$$v\mathcal{L}u - u\mathcal{L}^+ v = \frac{d}{dx}[u, v]$$

avec

$$[u, v] = \alpha_0 W(u, v) + (\alpha_1 - \alpha_0')uv$$

La vérification de cette identité est immédiate et on en déduit

[4] A l'aube du XIXe siècle, le mathématicien norvégien Niels Henrik Abel (1802–1829) allait révolutionner sa science, et Hermite a pu déclarer : "Il a laissé aux mathématiciens de quoi s'occuper pendant cinq cents ans". D'abord algébriste, il établit l'impossibilité de résolution par radicaux des équations algébriques de degré cinq et sa méthode ouvrait la voie aux travaux de Galois sur les groupes de substitution des racines d'une équation. A ce propos, il donne des critères de résolubilité par radicaux et étudie de nouveaux types d'équations, appelées de nos jours équations abéliennes, possédant cette propriété. En analyse, il est le fondateur, avec Jacobi, de la théorie des fonctions elliptiques.

[5] Joseph-Louis LAGRANGE (1736-1813), est issu d'une famille turinoise d'origine française, il a donné au calcul des variations sa formulation générale en l'abordant de manière purement analytique. Il appliquera ses méthodes à la mécanique, dont il donne un exposé systématique qui repose sur la théorie des équations différentielles.

Corollaire 5.2.4. (Formule de Green) *Pour toutes fonctions u et v dans $C^2(I)$ et x_1, x_2 dans I, on a*

$$\int_{x_1}^{x_2} (v\mathcal{L}u - u\mathcal{L}^+v)\,dx = [u,v](x_2) - [u,v](x_1)$$

La formule de Green est souvent utilisée lorsque le second membre est nul. Elle s'écrit de façon très simple lorsque l'opérateur $\mathcal{L}$ est formellement auto-adjoint

Définition 5.2.5. *L'opérateur différentiel $\mathcal{L}$ est dit formellement auto-adjoint si, pour tout u dans $C^2(I)$, $\mathcal{L}u = \mathcal{L}^+u$.*

Théorème 5.2.6. (a) *L'opérateur $\mathcal{L}$ est formellement auto-adjoint si, et seulement si, les coefficients α_0 et α_1 sont reliés par $\alpha_0' = \alpha_1$.*

(b) *Tout opérateur L, formellement auto-adjoint (à coefficients réels), s'écrit sous la forme*

$$Lu = (pu')' + qu$$

et pour un tel opérateur, la formule de Green s'écrit

$$\int_{x_2}^{x_2} (uLv - vLu)dy = p(x_2)W(u,v)(x_2) - p(x_1)W(u,v)(x_1)$$

(c) *Si $Lu = 0$ et $Lv = 0$, alors $pW(u,v)$ est une constante.*

Démonstration. (a) On vérifie facilement que pour toute fonction u dans $C^2(I)$

$$\mathcal{L}^+u = \alpha_0 u'' + (2\alpha_0' - \alpha_1)u' + (\alpha_0'' - \alpha_1' + \alpha_2)u$$

Il en résulte que $\mathcal{L} = \mathcal{L}^+$ si, et seulement si,

$$2\alpha_0' - \alpha_1 = \alpha_1 \quad \text{et} \quad \alpha_2 = \alpha_0'' - \alpha_1' + \alpha_2$$

ce qui est équivalent à la relation $\alpha_1 = \alpha_0'$. L'assertion (b) se déduit de (a), avec $\alpha_0 = p$ et $\alpha_2 = q$, et de la proposition 2.2 ; quant à l'assertion (c), elle découle de (b). □

L'assertion (b) montre l'avantage d'avoir un opérateur formellement auto-adjoint. Par exemple, lorsque u et v sont à support compact dans I et à valeurs complexes, la formule de Green s'écrit

$$\int_I (\overline{v}Lu - u\overline{Lv})dy = 0$$

c'est-à-dire $\langle Lu, v\rangle = \langle u, Lv\rangle$, où le produit scalaire est celui de l'espace $L^2(I, dx)$. Cela justifie le qualificatif "formellement auto-adjoint" qui devrait être précisé en ajoutant "dans $L^2(I, dx)$".

Le théorème qui suit dit précisément que tout opérateur différentiel du second ordre, linéaire et à coefficients réels, peut être transformé en un opérateur formellement auto-adjoint.

Théorème 5.2.7. *Soit $\mathcal{L}$ l'opérateur différentiel défini par*

$$\mathcal{L}u = \alpha_0 u'' + \alpha_1 u' + \alpha_2 u$$

Alors l'opérateur $\mu\mathcal{L}$, avec

$$\mu(x) = \exp\Big(\int_\tau^x \frac{\alpha_1 - \alpha_0'}{\alpha_0}\, dy\Big)$$

est formellement auto-adjoint (τ est quelconque dans I).

Démonstration. D'après le théorème précédent, l'opérateur μL est formellement auto-adjoint si, et seulement si, $(\mu\alpha_0)' = \mu\alpha_1$. L'expression de μ s'en déduit immédiatement. □

Exemple 5.2.8. -

(a) *L'opérateur de Legendre* est donné pour $-1 < x < 1$ par

$$Lu = (1 - x^2)u'' - 2xu' \quad \text{ou} \quad Lu = ((1 - x^2)u')'$$

c'est un opérateur formellement auto-adjoint dans $L^2(-1, 1)$.

(b) *L'opérateur de Laguerre* est donné pour $x \geq 0$ par

$$Lu = xu'' + (1 - x)u'$$

Le facteur μ est dans ce cas

$$\mu(x) = \exp\Big(\int_0^x \frac{(1-y) - 1}{y}\, dy\Big) = e^{-x}$$

L'opérateur $e^{-x}L$ est donc formellement auto-adjoint dans l'espa-ce $L^2(0, +\infty)$. Autrement dit, pour des fonctions f et g dans $C^2(0, \infty)$ et à supports compacts dans $(0, +\infty)$

$$\int_0^\infty e^{-x} Lf(x)\overline{g(x)}\, dx = \int_0^\infty f(x) e^{-x} \overline{Lg(x)}\, dx$$

Cela revient à dire que l'opérateur de Laguerre est formellement auto-adjoint dans $L^2((0, +\infty); e^{-x}dx)$, celui-ci est donc l'espace *naturel* dans lequel on étudie l'opérateur de Laguerre.

(c) *L'opérateur d'Hermite* est donné pour x réel par

$$Lu = u'' - xu'$$

Le facteur μ est, dans ce cas,

$$\mu(x) = \exp\Big(\int_0^x -y\,dy\Big) = e^{-x^2/2}$$

et par suite $e^{-x^2/2}L$ est un opérateur formellement auto-adjoint dans l'espace $L^2(\mathbb{R})$. Pour deux fonctions f et g dans $C^2(\mathbb{R})$ et à supports compacts

$$\int_{\mathbb{R}} e^{-x^2/2} Lf(x)\overline{g(x)}\,dx = \int_{\mathbb{R}} f(x)e^{-x^2/2}\overline{Lg(x)}\,dx$$

Cette égalité traduit le fait que l'opérateur d'Hermite est formellement auto-adjoint dans $L^2(\mathbb{R}; e^{-x^2/2}dx)$ qui est donc l'espace *adapté* à l'étude de cet opérateur.

EXERCICES

1. Soient u_1 et u_2 deux solutions linéairement indépendantes de l'équation homogène $\mathcal{L}u = 0$. Montrer que leur wronskien W vérifie l'équation différentielle

$$W'(u_1, u_2) = -\frac{\alpha_1}{\alpha_0}W(u_1, u_2)$$

En déduire la formule (1) dite formule d'Abel.
Solution : Par définition, $W(u_1, u_2) = u_1u_2' - u_2u_1'$, par dérivation on obtient $W'(u_1, u_2) = u_1u_2'' - u_2u_1''$. Comme $\mathcal{L}u_1 = \mathcal{L}u_2 = 0$, on obtient l'équation $W'(u_1, u_2) = -(\alpha_1/\alpha_0)W(u_1, u_2)$. En intégrant cette équation différentielle, on obtient la formule d'Abel. □

2. On suppose que dans l'expression de $\mathcal{L}$, $\alpha_0 = 1$. Soient u_1 et u_2 deux solutions linéairement indépendantes de l'équation $\mathcal{L}u = 0$. Montrer que pour toute fonction f continue, la fonction

$$v(x) = \int_{x_0}^x \frac{u_2(x)u_1(t) - u_1(x)u_2(t)}{W(u_1, u_2)(t)} f(t)\,dt$$

est une solution particulière de l'équation différentielle $\mathcal{L}u = f$. (v se calcule à l'aide de la méthode de la variation des constantes).

3. On considère l'équation de Legendre (voir chapitre II, section 4)

$$(1 - x^2)u'' - 2xu' + n(n+1)u = 0, \quad n \in \mathbb{N}$$

On sait que le polynôme de Legendre P_n est une solution de cette équation vérifiant $P_n(1) = 1$. En utilisant la formule d'Abel, montrer

qu'une deuxième solution de l'équation de Legendre, linéairement indépendante de P_n, est fournie par

$$Q_n(x) = P_n(x) \int \frac{dt}{(1-t^2)P_n(t)^2}$$

Calculer Q_0 et Q_1. (les Q_n sont appelés fonctions de Legendre de deuxième espèce).
Solution. L'expression de Q_n est une conséquence directe de la formule (1) et du fait qu'une primitive de $(2t)/(1-t^2)$ est $-\mathrm{Log}(1-t^2)$. D'autre part, comme $P_0 = 1$ et $P_1(x) = x$, la formule (1) donne $Q_0(x) = (1/2)\mathrm{Log}[(1+x)(1-x)^{-1}]$ et $Q_1(x) = xQ_0(x) - 1$. □

4. L'équation d'Hermite (chapitre II, section 5), $u'' - xu' + nu = 0$ admet, pour chaque entier n, le polynôme d'Hermite H_n comme solution. Montrer que la fonction

$$h_n(x) = H_n(x) \int H_n^{-2}(t)e^{t^2/2}\, dt$$

est une deuxième solution linéairement indépendante de H_n. (usuellement, on appelle fonction d'Hermite de seconde espèce la fonction $(n!)h_n$). Montrer que si l'on ajoute la condition $h_n(0) = 0$, alors la fonction h_0 est donnée par $h_0(x) = \int_0^x e^{t^2/2}\, dt$.

5. L'équation de Laguerre $xu'' + (1-x)u' + nu = 0$ admet, pour chaque entier n, le polynôme de Laguerre L_n comme solution, (voir chapitre II, section 6). Montrer que la fonction définie par

$$\ell_n(x) = L_n(x) \int t^{-1} L_n^{-2}(t)e^t\, dt$$

est une deuxième solution linéairement indépendante de L_n.

6. Dans chacun des exemples suivants, on donne une solution u_1 de l'équation différentielle et on demande de construire une deuxième solution linéairement indépendante sur l'intervalle indiqué
(a) $u'' - (2/x^2)u = 0$, $u_1(x) = x^2$, $0 < x < \infty$.
(b) $u'' - 4xu' + (4x^2 - 2)u = 0$, $u_1(x) = e^{x^2}$, $0 < x < \infty$.
(c) $(1-x^2)u'' - 2xu' + 2u = 0$, $u_1(x) = x$, $0 < x < 1$.
Solution : Il suffit d'appliquer la formule (2).

7. Dans les exemples qui suivent, utiliser la solution u_1 proposée de l'équation homogène pour trouver la solution générale de l'équation avec second membre sur l'intervalle indiqué.
(a) $u'' - (2/x^2)u = e^x$, $u_1(x) = x^2$, $0 < x < \infty$.
(b) $(1-x^2)u'' - 2xu' + 2u = f(x)$ $u_1(x) = x$, $0 < x < 1$
où f est une fonction continue sur $[0, 1]$.
(c) $x^2u'' + xu' + (x^2 - 1/4)u = 3\sqrt{x}\sin x$, $u_1(x) = \sqrt{x}\sin x$, avec $0 < x < \infty$.

Solution : Grâce à la formule (2), on cherche d'abord une deuxième solution u_2 linéairement indépendante de u_1. Ensuite, on utilise la méthode de la variation des constantes (exercice 2) pour trouver une solution particulière. □

8. Soit $\mathcal{L}$ l'opérateur différentiel défini par

$$\mathcal{L}u = (1-x^2)u'' - xy' + \alpha u, \quad -1 < x < 1$$

(a) Calculer l'adjoint formel de $\mathcal{L}$.
(b) Transformer $\mathcal{L}$ en un opérateur formellement auto-adjoint dans l'espace $L^2(-1,1)$.

6. Transformer l'opérateur hypergéométrique défini par

$$x(x-1)u'' + [(1+\alpha+\beta)x - \gamma]u' + \alpha\beta u$$

où α, β et γ sont des constantes réelles, en un opérateur formellement auto-adjoint dans $L^2(-1,1)$.

9. Transformer chacun des opérateurs différentiels qui suivent, en un opérateur formellement auto-adjoint :
(a) $\mathcal{L}(u) = u'' - 2xu'$
(b) $\mathcal{L}u = x^2u'' + xu' + x^2u$.

10. Soit $\mathcal{L}$ l'opérateur différentiel défini par $\mathcal{L}u = \alpha_0 u'' + \alpha_1 u' + \alpha_2 u$, où α_0, α_1 et α_2 sont des fonctions données sur un intervalle I à valeurs réelles.
(a) Déterminer les fonctions p, q et r telles que

$$\mathcal{L}u = \frac{1}{r(x)}\left\{\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u\right\}$$

(b) Montrer que le changement de variables $t = \int^x \sqrt{r(s)/p(s)}ds$ transforme le deuxième membre de l'égalité ci-dessus sous la forme

$$\mathcal{L}u = \tilde{u}''(t) + \frac{\beta'(t)}{2\beta(t)}\tilde{u}'(t) + \frac{\tilde{q}(t)}{\tilde{r}(t)}\tilde{u}(t), \quad \text{avec} \quad \tilde{u}(t) = u(x)$$

où l'on a posé $\tilde{u}(t) = u(x)$ et $\tilde{r}(t) = r(x)$, $\tilde{q}(t) = q(x)$, $\tilde{p}(t) = p(x)$ et $\beta(t) = \tilde{r}(t)\tilde{p}(t)$.
(c) Montrer que si l'on pose $\tilde{u}(t) = \beta(t)^{-1/4}v(t)$, le membre de droite de l'expression ci-dessus devient

$$\mathcal{L}u(x) = \beta(t)^{-1/4}[v''(t) - Q(t)v(t)]$$

où la fonction Q est à déterminer. En déduire que la "transformation de Liouville"

$$t = \int^x \sqrt{r(s)/p(s)}ds, \quad u = ((p(x)r(x))^{-1/4}v$$

permet de passer de l'équation

$$\frac{1}{r(x)}\left\{\frac{d}{dx}\left(p(x)\frac{du}{dx}\right)+q(x)u\right\}=\lambda u$$

à l'équation

$$v''(t)-Q(t)v(t)=\lambda v(t)$$

Noter que l'opérateur différentiel qui apparaît dans la premiere équation est formellement auto-adjoint dans $L^2(I,r(x)dx)$, alors que celui intervenant dans la deuxième est formellement auto-adjoint dans $L^2(dt)$.

Solution : (a) On vérifie que les fonctions p, q et r sont données par : $r(x)=\alpha_0(x)/p(x)$,

$$p(x)=\exp\left(\int^x\frac{\alpha_1(s)}{\alpha_0(s)}ds\right),\quad \text{et}\quad q(x)=\frac{\alpha_2(x)p(x)}{\alpha_0(x)}$$

(c) La fonction Q est donnée par

$$Q(t)=\frac{\beta''}{4\beta}-3\frac{\beta'}{4\beta}-\frac{\tilde{q}}{\tilde{r}}$$

5.3 Opérateur de Sturm-Liouville Régulier

Un opérateur de Sturm[6]-Liouville[7] est la donnée d'un opérateur différentiel linéaire et du second ordre (qu'on peut supposer, d'après le théorème 2.6, formellement auto-adjoint)

$$Lu=\frac{d}{dx}\left(p(x)\frac{du}{dx}\right)-q(x)u$$

sur un intervalle $[a,b]$ et de conditions aux extrémités de l'intervalle, appelées conditions au bord (ou à la frontière)

$$p(a)u'(a)\sin\theta-u(a)\cos\theta=0$$
$$p(b)u'(b)\sin\gamma-u(b)\cos\gamma=0$$

[6]C'est à partir de 1830 que le mathématicien français Charles François STURM (1803-1855), en liaison avec son ami Liouville, aborde le problème de la théorie générale des oscillations et étudie les équations différentielles du second ordre. Les méthodes employées seront à l'origine de nombreux travaux et découvertes mathématiques.

[7]Le mathématicien français Joseph LIOUVILLE (1809-1882) est le fondateur du Journal de Mathématiques Pures et Appliquées, appelé traditionellement Journal de Liouville. Les deux premiers volumes (1836-1837) contiennent six mémoires, les uns de Liouville, les autres de Sturm, sur le problème qui porte aujourd'hui leurs noms.

La fonction p est dans C^1 et strictement positive sur $[a, b]$, la fonction q est réelle et continue sur $[a, b]$.

Le domainde D_L de l'opérateur de Sturm-Liouville est l'espace des fonctions dans $C^2[a, b]$ vérifiant les conditions à la frontière. L'opérateur de Sturm-Liouville, associé à l'opérateur différentiel et aux conditions à la frontière, est l'application

$$L : D_L \to C[a, b]$$

et on parlera alors de l'opérateur (D_L, L). Le problème de Sturm-Liouville consiste à résoudre le système

$$\begin{cases} \lambda u - Lu = f \\ u \in D_L \end{cases} \tag{1}$$

où f est une fonction continue sur $[a, b]$ donnée, λ est un paramètre complexe et où u est la fonction inconnue à chercher.

Soient u et v deux fonctions dans $C^2[a, b]$. On désigne par $W(u, v)$ leur wronskien c'est-à-dire $W(u, v) = uv' - vu'$ et on pose

$$[u, v] = pW(u, v)$$

On a

$$\frac{d}{dx}[u, v] = uLv - vLu(1)$$

de telle sorte que

$$\int_a^b (uLv - vLu)\, dx = [u, v](b) - [u, v](a)$$

et si u et v sont dans D_L, le second membre de l'égalité ci-dessus est nul et on obtient

$$\int_a^b uLv\, dx = \int_a^b vLu\, dx \tag{5.6}$$

Pour cette raison, on dit que l'opérateur (D_L, L) est symétrique ou formellement auto-adjoint.

EXERCICES

1. Soit (D_L, L) l'opérateur défini par : $Lu = u''$ et D_L l'ensemble des fonctions u dans $C^2[0, \pi]$ vérifiant les conditions au bord

$$\begin{cases} u(\pi) = au(0) + bu'(0) \\ u'(\pi) = cu(0) + du'(0) \end{cases}$$

Expliquer comment choisir les réels a, b, c et d pour que (D_L, L) soit symétrique, c'est-à-dire pour que l'on ait

$$\langle Lu, v\rangle = \langle u, Lv\rangle, \quad \forall\, u, v \in D_L$$

Solution. En posant $W(u,v) = uv' - vu'$, on vérifie que
$$\langle Lu, v\rangle - \langle u, Lv\rangle = W(u,v)(0) - W(u,v)(\pi)$$
L'opérateur (D_L, L) est donc symétrique si, et seulement si, on a $W(u,v)(\pi) = W(u,v)(0)$, pour tout u et tout v dans le domaine D_L. Or, pour de tels éléments, $W(u,v)(\pi) = (ad - bc)W(u,v)(0)$ et la condition précédente est équivalente à : $ad - bc = 1$. □

2. Soient α et β deux nombres complexes et posons $Lu = u''$, pour u dans $C^2[a,b]$ à valeurs complexes. Le produit scalaire est défini par $\langle f, g\rangle = \int_a^b f\overline{g}\,dx$.
(i) Soit $D_1 = \{\, u \in C^2[a,b] \mid u'(a) = \alpha u(a), u'(b) = \beta u(b)\,\}$. Montrer que si α et β sont réels, alors (D_1, L) est symétrique.
(ii) Soit $D_2 = \{\, u \in C^2[a,b] \mid u(a) = \alpha u'(a), u(b) = \beta u'(b)\,\}$. Montrer que si α et β sont réels, alors (D_2, L) est symétrique.
(iii) Soit $D_3 = \{\, u \in C^2[a,b] \mid u(b) = \alpha u(a), u'(b) = \beta u'(a)\,\}$. Montrer que si $\alpha\overline{\beta} = 1$, alors (D_3, L) est symétrique.
(iv) Soit $D_4 = \{\, u \in C^2[a,b] \mid u'(a) = \alpha u(b), u'(b) = \beta u(a)\,\}$. Montrer que si $\alpha + \overline{\beta} = 0$, alors (D_4, L) est symétrique.
Solution : Dans chacun des cas, il s'agit de calculer $W(u,\overline{v})$ pour u et v dans D_j, $1 \le j \le 4$ et de demander à ce que l'on ait

$$W(u,\overline{v})(a) - W(u,\overline{v})(b) = 0 \qquad (*)$$

Par exemple, dans le cas (iv) le premier membre de $(\star)$ vaut $-(\alpha + \overline{\beta})u(b)\overline{v}(a) + (\overline{\alpha} + \beta)u(a)\overline{v(b)}$, il est nul si $\overline{\alpha} + \beta = 0$. Dans le cas (iii), on trouve $W(u,\overline{v})(a) = u(a)\overline{v}'(a) - u'(a)\overline{v}(a)$ et $W(u,\overline{v})(b) = \alpha\overline{\beta}u(a)\overline{v}'(a) - \overline{\alpha}\beta u'(a)\overline{v}(a)$. On en déduit que si $\alpha\overline{\beta} = 1$, la condition $(\star)$ est satisfaite. □

5.4 Fonction de Green et Résolvante

Soit u_1 la solution de $Lu = 0$ qui vérifie

$$u_1(a) = \sin\theta, \quad p(a)u_1'(a) = \cos\theta$$

et soit u_2 la solution de $Lu = 0$ qui vérifie

$$u_2(b) = \sin\gamma, \quad p(b)u_2'(b) = \cos\gamma$$

La relation (1) montre que la fonction $[u_1, u_2]$ est constante. Comme p est strictement positive sur $[a, b]$, cette constante est nulle si, et seulement si, le wronskien $W(u_1, u_2)$ est identiquement nul, ce qui équivaut à dire que u_1 et u_2 sont liées. Dans ce cas u_1 est dans D_L et par suite l'opérateur (D_L, L) est non injectif. En revanche, si l'opérateur (D_L, L) est supposé injectif, ce qui précède montre que les solutions u_1 et u_2 sont nécessairement liéairement indépendantes.
On suppose que l'opérateur (D_L, L) est injectif, c'est-à-dire que le problème

$$\begin{cases} Lu = 0 \\ u \in D_L \end{cases}$$

n'admet que la solution nulle $u = 0$. Les fonctions u_1 et u_2, sont donc linéairement indépendantes.

On va résoudre, par la méthode de la variation des constantes de Lagrange, le problème

$$\begin{cases} Lu = -f \\ u \in D_L \end{cases} \tag{II}$$

où f est une fonction donnée, continue sur $[a, b]$. La méthode consiste à poser

$$u(x) = c_1(x)u_1(x) + c_2(x)u_2(x)$$

où les fonctions c_1 et c_2 vérifient $u_1c_1' + u_2c_2' = 0$, de telle sorte que l'on ait

$$\begin{cases} p(u_1'c_1' + u_2'c_2') = f \\ u_1c_1' + u_2c_2' = 0 \end{cases}$$

d'où il vient $[u_1, u_2]c_1' = u_2f$ et $[u_1, u_2]c_2' = -u_1f$. En tenant compte du fait que u appartient à D_L, c'est-à-dire satisfait les conditions à la frontière, on trouve $c_2(a) = 0$ et $c_1(b) = 0$ et par suite l'intégration du sytème précédent donne

$$c_1(x) = -\int_x^b \frac{u_2(y)f(y)}{[u_1, u_2]}\,dy, \quad c_2(x) = -\int_a^x \frac{u_1(y)f(y)}{[u_1, u_2]}\,dy$$

Ainsi, le problème (II) admet une solution qui s'écrit

$$u(x) = \int_a^b G(x, y)f(y)\,dy$$

avec

$$G(x, y) = \begin{cases} \dfrac{-1}{[u_1, u_2]}u_1(y)u_2(x), & \text{si } a \leq y \leq x \leq b\,; \\ \dfrac{-1}{[u_1, u_2]}u_1(x)u_2(y), & \text{si } a \leq x \leq y \leq b. \end{cases} \tag{1}$$

On peut donc énoncer

Théorème 5.4.1. *On suppose l'opérateur (D_L, L) injectif.*

(i) *Soient f une fonction de $C[a,b]$ et u la fonction définie par*

$$u(x) = \int_a^b G(x,y) f(y)\, dy$$

alors, la fonction u appartient à D_L et vérifie $Lu = -f$

(ii) *Soient u une fonction de D_L et $f = -Lu$, alors*

$$u(x) = \int_a^b G(x,y) f(y)\, dy$$

Démonstration. L'assertion (i) est déjà démontrée. Posons

$$v(x) = \int_a^b G(x,y) f(y)\, dy$$

d'après l'assertion (i) la fonction v appartient à D_L et $Lv = -f$, donc $L(u - v) = 0$. Comme l'opérateur (D_L, L) est injectif, on en déduit que $u = v$. □

Définition 5.4.2. *La fonction $G(.,.)$ s'appelle la fonction (ou le noyau) de Green*[8] *de l'opérateur (D_L, L).*

Le théorème 4.1 exprime que si l'opérateur $(D_L, -L)$ est injectif, alors il est inversible et que son inverse est l'opérateur intégral à noyau donné par

$$Gf(x) = \int_a^b G(x,y) f(y)\, dy$$

La fonction de Green est caractérisée par les propriétés suivantes qui permettent donc de la construire.

Proposition 5.4.3. *Soit G_x la fonction $y \mapsto G(x,y)$.*

(a) *La fonction G_x vérifie les conditions à la frontière en a et en b, elle est ce classe C^2 sur $[a, x[$ et $]x, b]$ et sur chacun de ces intervalles elle satisfait : $LG_x = 0$*

(b) *En x la dérivée de G_x est discontinue et le saut en ce point est*

$$\frac{d}{dy} G_x(x+0) - \frac{d}{dy} G_x(x-0) = -\frac{1}{p(x)}$$

Si δ_x désigne la masse de Dirac au point x, ces propriétés montrent qu'au sens des distributions, $LG_x = -\delta_x$.

[8] George GREEN (1793-1841), est un mathématicien anglais qui, à travers sa recherche d'une formulation de la théorie de l'électricité statique et du magnétisme, est le créateur de la théorie du potentiel. Boulanger de sa profession, il s'initia seul aux mathématiques, principalement en lisant les mémoires de Poisson.

Exemple 5.4.4. - Soit (D_L, L) l'opérateur défini par

$$D_L = \{ \ u \in C^2[0,1] \mid u(0) = 0, u(1) = 0 \ \}$$
$$Lu = u'', \ u \in D_L$$

Pour trouver son *inverse*, on doit résoudre

$$\begin{cases} u'' = -f \\ u(0) = 0, \ u(1) = 0 \end{cases}$$

Après deux intégrations par parties, on trouve

$$u(x) = -\int_0^x \left(\int_0^t f(y)dy \right) dt + c_1 x + c_2$$

où c_1 et c_2 sont deux constantes arbitraires. Intervertissant l'ordre des intégrations, on obtient

$$u(x) = -\int_0^x (x-y) f(y)dy + c_1 x + c_2$$

Ecrivant que $u(0) = u(1) = 0$, il vient

$$c_2 = 0 \quad \text{et} \quad c_1 = \int_0^1 (1-y) f(y)dy$$

et la solution u s'écrit donc

$$\begin{aligned} u(x) &= \int_0^x (1-x) y f(y)dy + \int_x^1 (1-y) x f(y)dy \\ &= \int_0^1 G(x,y) f(y)dy \end{aligned}$$

où l'on a posé

$$G(x,y) = \begin{cases} (1-x)y, & 0 \leq y \leq x \leq 1\,; \\ (1-y)x, & 0 \leq x \leq y \leq 1. \end{cases}$$

On vérifie facilement que G, ainsi trouvée, est la fonction de Green de l'opérateur (D_L, L) ; on peut la retrouver en appliquant la formule (1).

Exemple 5.4.5. - Soit (D_L, L) l'opérateur défini par

$$D_L = \{ \ u \in C^2[0,\pi]/u(0) = 0, u(\pi) = 0 \ \}$$
$$Lu = e^{-2x}[(e^{2x}u')' + e^{2x}u], \ u \in D_L$$

Pour inverser cet opérateur, on doit résoudre

$$\begin{cases} Lu = -f \\ u(0) = 0, \ u(\pi) = 0 \end{cases}$$

L'équation différentielle s'écrit

$$u'' + 2u' + u = -f$$

En posant $u = e^{-x}v$, l'équation devient $e^{-x}v'' = -f$, et sa solution générale est donc de la forme

$$v(x) = -\int_0^x (x-y)e^y f(y)dy + c_1 x + c_2$$

soit

$$u(x) = -\int_0^x (x-y)e^{(y-x)} f(y)dy + c_1 x e^{-x} + c_2 e^{-x}$$

Ecrivant que u vérifie les conditions aux bords en 0 et en π, on trouve

$$c_2 = 0 \quad \text{et} \quad c_1 = \frac{1}{\pi}\int_0^\pi (\pi - y)e^y f(y)dy$$

La solution u est donc donnée par

$$u(x) = \int_0^\pi G(x,y)e^{2y} f(y)dy$$

avec

$$G(x,y) = \begin{cases} (1/\pi)(\pi - x)ye^{-(x+y)}, & 0 \le y \le x \le \pi \\ (1/\pi)(\pi - y)xe^{-(x+y)}, & 0 \le x \le y \le \pi \end{cases}$$

On vérifie que la fonction G, ainsi trouvée, est bien la fonction de Green de l'opérateur (D_L, L).

Le théorème 4.1 s'applique à l'opérateur $(D_L, \lambda I - L)$, pourvu que celui-ci soit injectif. Dans ce cas, $(D_L, \lambda I - L)$ est inversible et son inverse, qu'on notera G_λ est de la forme

$$G_\lambda f(x) = \int_a^b G_\lambda(x,y)f(y)\,dy$$

où la fonction G_λ est construite de façon analogue au cas où $\lambda = 0$. Plus précisément, soit $u_1(.,\lambda)$ la solution de $\lambda u - Lu = 0$ vérifiant

$$u_1(a,\lambda) = \sin\theta, \quad p(a)u_1'(a,\lambda) = \cos\theta$$

et soit $u_2(.,\lambda)$ la solution de $\lambda u - Lu = 0$ vérifiant

$$u_2(b,\lambda) = \sin\gamma, \quad p(b)u_2'(b,\lambda) = \cos\gamma$$

Si $\lambda I - L$ est injectif, les solutions $u_1(.,\lambda)$ et $u_2(.,\lambda)$ sont linéairement indépendantes et par suite $[u_1(.,\lambda), u_2(.,\lambda)]$ est une constante non nulle qui ne dépend que de λ. On vérifie alors que le noyau G_λ est donné par

$$G_\lambda(x,y) = \begin{cases} \dfrac{-1}{[u_1(.,\lambda), u_2(.,\lambda)]} u_1(y,\lambda)u_2(x,\lambda), & \text{si } a \leq y \leq x \leq b\,; \\ \dfrac{-1}{[u_1(.,\lambda), u_2(.,\lambda)]} u_1(x,\lambda)u_2(y,\lambda), & \text{si } a \leq x \leq y \leq b. \end{cases} \quad (2)$$

le théorème 4.1 se traduit par

Théorème 5.4.6. *On suppose l'opérateur* $(D_L, \lambda I - L)$ *injectif*

(i) *Soient* f *une fonction continue sur* $[a,b]$ *et* u *la fonction définie par*

$$u(x) = \int_a^b G_\lambda(x,y) f(y)\,dy$$

Alors, la fonction u *appartient à* D_L *et vérifie* $\lambda u - Lu = f$

(ii) *Soient* u *une fonction de* D_L *et* $f = \lambda u - Lu$. *alors*

$$u(x) = \int_a^b G_\lambda(x,y) f(y)\,dy$$

La famille des opérateurs G_λ s'appelle la résolvante de l'opérateur (D_L, L).

Dans la suite on pose

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx, \quad \text{et} \quad \|f\| = \sqrt{\langle f, f\rangle}$$

L'espace $C[a,b]$ est ainsi muni d'une structure d'espace préhilbertien dont le complété est l'espace $L^2[a,b]$.

Définition 5.4.7. *Un nombre complexe* λ *est une valeur propre de* (D_L, L) *s'il existe une fonction* u *dans* D_L, *non nulle et vérifiant*

$$\lambda u - Lu = 0$$

la fonction u *est alors appelée une fonction propre de* (D_L, L) *associée à la valeur propre* λ.

Théorème 5.4.8. *Les valeurs propres de l'opérateur* (D_L, L) *sont réelles. Les sous-espaces propres correspondant sont de dimension 1 et deux à deux orthogonaux.*

Démonstration. La formule (1) du paragraphe précédent se traduit par $\langle Lu, v\rangle = \langle u, Lv\rangle$ si u et v sont dans D_L, de sorte que si u appartient à D_L

le nombre $\langle Lu, u\rangle$ est réel. Soient λ une valeur propre de (D_L, L) et ϕ une fonction propre associée

$$\langle L\phi, \phi\rangle = \lambda\|\phi\|^2$$

la valeur propre λ est donc réelle. Si ψ est une autre fonction propre associée à λ. La formule (1) montre que $pW(\phi, \psi)$ est une constante, comme sa valeur en a est nulle, on a $W(\phi, \psi) = 0$ et les fonctions ϕ et ψ sont donc proportionnelles.

Soient ϕ et ψ deux fonctions propres correspondant aux valeurs propres λ et ν

$$\langle L\phi, \psi\rangle = \lambda\langle \phi, \psi\rangle = \langle \phi, L\psi\rangle = \nu\langle \phi, \psi\rangle$$

d'où $(\lambda - \nu)\langle \phi, \psi\rangle = 0$ et si $\lambda \neq \nu$, $\langle \phi, \psi\rangle = 0$. □

Exemple 5.4.9. - Soit (D_L, L) l'opérateur défini par

$$D_L = \{\ u \in C^2[0, 1] \mid u(0) = 0, u(1) = 0\ \}$$
$$Lu = u'', \ u \in D_L$$

on a, avec les notations utilisées,

$$u_1(x, \lambda) = \frac{\operatorname{sh}\sqrt{\lambda}x}{\sqrt{\lambda}}, \quad u_2(x, \lambda) = \frac{\operatorname{sh}\sqrt{\lambda}(x-1)}{\sqrt{\lambda}}, \quad [u_1, u_2] = \frac{\operatorname{sh}\sqrt{\lambda}}{\sqrt{\lambda}}$$

Les valeurs propres de l'opérateur (D_L, L) sont les complexes λ tels que $[u_1, u_2] = 0$, on en déduit que les valeurs propres de (D_L, L) et les fonctions propres (normalisées) associées sont

$$\lambda_n = -n^2\pi^2, \quad \phi_n(x) = \sqrt{2}\sin(n\pi x), \quad n \geq 1$$

Compte tenu de la formule (2), la fonction de Green G_λ est définie pour $\lambda \neq \lambda_n$ par

$$G_\lambda(x, y) = \frac{\sqrt{\lambda}}{\operatorname{sh}\sqrt{\lambda}} \begin{cases} \dfrac{\operatorname{sh}\sqrt{\lambda}x}{\sqrt{\lambda}} \dfrac{\operatorname{sh}\sqrt{\lambda}(1-y)}{\sqrt{\lambda}}, & 0 \leq x \leq y \leq 1 \\ \dfrac{\operatorname{sh}\sqrt{\lambda}y}{\sqrt{\lambda}} \dfrac{\operatorname{sh}\sqrt{\lambda}(1-x)}{\sqrt{\lambda}}, & 0 \leq y \leq x \leq 1 \end{cases}$$

Exemple 5.4.10. - Soit (D_L, L) l'opérateur défini par

$$D_L = \{\ u \in C^2[0, \pi] \mid u(0) = 0, u(\pi) = 0\ \}$$
$$Lu = e^{-2x}[(e^{2x}u')' + e^{2x}u], \quad u \in D_L$$

On vérifie que ses valeurs propres et les fonctions propres associées sont

$$\lambda_n = -n^2, \quad \phi_n(x) = c_n e^{-x}\sin nx, \quad n \geq 1$$

où c_n est choisie de façon que ϕ_n soit normalisée. On calculera c_n et on explicitera l'expression de la fonction de Green G_λ, $\lambda \neq \lambda_n$, en exercice.

EXERCICES

1. Montrer que si λ et μ ne sont pas des valeurs propres de l'opérateur (D_L, L), G_λ vérifie l'équation résolvante

$$G_\lambda - G_\mu = (\mu - \lambda) G_\lambda G_\mu$$

Montrer que cette équation se traduit par

$$G_\lambda(x, y) - G_\mu(x, y) = (\mu - \lambda) \int_a^b G_\lambda(x, z) G_\mu(z, y)\, dz$$

Montrer que pour tout f dans $L^2[a, b]$

$$\|G_\lambda f\| \leq \frac{1}{|\Im\mathrm{m}\lambda|} \|f\|$$

Solution : Montrons cette dernière inégalité. Posons $\lambda = \xi + i\eta$ et $u = G_\lambda f$, nous avons $f = \xi u + i\eta u - Lu$ et donc

$$\begin{aligned} \|f\|^2 &= \langle \xi u - Lu, \xi u - Lu \rangle + i\eta \langle u, \xi u - Lu \rangle + \eta^2 \|u\|^2 \\ &= \|\xi u - Lu\|^2 + \eta^2 \|u\|^2 \geq \eta^2 \|u\|^2 \end{aligned}$$

d'où on déduit que $\|G_\lambda f\|^2 \leq \eta^{-2} \|f\|^2$.

2. Montrer que si λ_0 n'est pas valeur propre de (D_L, L), le problème

$$\begin{cases} \lambda u - Lu = f \\ u \in D_L \end{cases}$$

est équivalent à l'équation intégrale

$$(\lambda - \lambda_0) G_{\lambda_0} u + u = G_{\lambda_0} f$$

En déduire que (D_L, L) et l'opérateur intégral G_{λ_0} admettent les mêmes sous-espaces propres et que λ est valeur propre de (D_L, L) si, et seulement si, $\mu = (\lambda_0 - \lambda)^{-1}$ est valeur propre de G_{λ_0}.

3. On considère l'opérateur (D_L, L) défini par

$$D_L = \{\ u \in C^2[0, \pi] \mid u(0) = 0, u(\pi) = 0\ \}$$
$$Lu = u'', u \in D_L$$

(a) Montrer que (D_L, L) est injectif et trouver, pour toute f continue sur $[0, \pi]$, la solution du problème

$$\begin{cases} Lu = -f \\ u \in D_L \end{cases}$$

(b) En déduire la fonction de Green $G(.,.)$ de (D_L, L).

(c) Déterminer les valeurs propres (λ_n) de (D_L, L) et les fonctions propres associées.
(d) Pour $\lambda \neq \lambda_n$ et f continue, trouver la solution du problème

$$\begin{cases} \lambda u - Lu = f \\ u \qquad\qquad \in D_L \end{cases}$$

(e) En déduire la fonction de Green $G_\lambda(.,.)$ et montrer que

$$\lim_{\lambda \to 0} G_\lambda(x, y) = G(x, y)$$

Solution : Il suffit de reprendre l'exemple 4.9 dans lequel on remplace l'intervalle $[0, 1]$ par $[0, \pi]$. D'abord, la solution générale de $Lu = 0$ étant de la forme $u(x) = c_1 + c_2 x$, la seule solution appartenant à D_L est 0 ; l'opérateur (D_L, L) est donc injectif. La méthode de la variation des constantes montre que toute solution de $Lu = -f$ s'écrit $u = c_1 + c_2 x + \int_0^x (t - x) f(t)\, dt$. La condition $u(0) = 0$ implique $c_1 = 0$, la condition $u(\pi) = 0$ donne $c_2 = (1/\pi) \int_0^\pi (\pi - t) f(t)\, dt$ et par suite la solution u de l'équation $Lu = -f$, qui appatient à D_L s'écrit $u = Gf$, où G est l'opérateur intégral de noyau la fonction de Green $G(.,.)$ donnée par

$$G(x, t) = \begin{cases} \dfrac{x(\pi - t)}{\pi}, & \text{si } 0 \leq x \leq t \leq \pi\,; \\ \dfrac{t(\pi - x)}{\pi}, & \text{si } 0 \leq t \leq x \leq \pi. \end{cases}$$

On remarque que l'expression de G peut être déduite de la formule (2). Nous avons répondu aux questions (a) et (b). Comme dans l'exemple 4.9, on montre que les valeurs propres (λ_n) et les fonctions propres (normalisées) sont données par

$$\lambda_n = -n^2, \quad \phi_n(x) = \sqrt{2/\pi} \sin(nx), \quad n \geq 1$$

De même, pour $\lambda \neq \lambda_n$, $\forall n \geq 1$, la solution $u \in D_L$ de l'équation $\lambda u - Lu = f$ est donnée par $u = G_\lambda f$, où G_λ est l'opérateur intégral de noyau la fonction de Green $G_\lambda(.,.)$ définie par

$$G_\lambda(x, t) = \begin{cases} \dfrac{\operatorname{sh} \sqrt{\lambda} x \operatorname{sh} \sqrt{\lambda}(\pi - t)}{\sqrt{\lambda} \operatorname{sh} \sqrt{\lambda} \pi}, & \text{si } 0 \leq x \leq t \leq \pi\,; \\ \dfrac{\operatorname{sh} \sqrt{\lambda} t \operatorname{sh} \sqrt{\lambda}(\pi - x)}{\sqrt{\lambda} \operatorname{sh} \sqrt{\lambda} \pi}, & \text{si } 0 \leq t \leq x \leq \pi. \end{cases}$$

La question (e) s'en déduit immédiatement. □

4. L'opérateur (D_L, L) étant celui défini dans l'exercice 3, résoudre le problème suivant où k est un réel donné

$$\begin{cases} \lambda u - Lu = \cos(kx) \\ u \in D_L \end{cases}$$

Solution : Si $\lambda \neq -k^2$, une solution particulière est donnée par la fonction $\cos(kx)/(\lambda + k^2)$. La solution générale de $Lu - \lambda u = 0$ est donc

$$u(x) = c_1 \operatorname{ch}(\sqrt{\lambda}x) + c_2 \frac{\operatorname{sh}(\sqrt{\lambda}x)}{\sqrt{\lambda}} + \frac{\cos(kx)}{\lambda + k^2}$$

Il suffit alors de choisir les constantes c_1 et c_2 de façon que $u(0) = u(\pi) = 0$, ce qui donne

$$u(x) = \frac{\operatorname{ch}(\sqrt{\lambda}x)}{\lambda + k^2} + \frac{\cos(kx)}{\lambda + k^2} + \frac{\operatorname{sh}(\sqrt{\lambda}x)}{(\lambda + k^2)\operatorname{sh}\sqrt{\lambda}\pi}(\operatorname{ch} k\pi - \operatorname{ch}(\sqrt{\lambda}\pi))$$

Si $\lambda = -k^2$, c'est-à-dire si λ est une valeur propre de (D_L, L), on construit une solution particulière u_0 de $\mathcal{L}u + k^2 u = \cos kx$ à partir des solutions linéairement indépendantes, $u_1(x) = \sin kx$ et $u_2(x) = \cos kx$, de $\mathcal{L}u + k^2 u = 0$, en utilisant la méthode de la variation des constantes (voir l'exercice 2 du paragraphe 2) ; il vient $u_0(x) = (1/k)\int_0^x \cos kt \sin k(x-t)\,dt = x \sin kx/(2k)$. La solution de notre problème est donc

$$u(x) = c_2 \frac{\sin kx}{k} + \frac{x \sin kx}{2k} \quad \text{où } c_2 \text{ est une constante arbitraire}$$

5. Déterminer les fonctions propres normalisées et l'expression de la fonction de Green G_λ de l'opérateur défini dans l'exemple 4.10.

6. On considère l'opérateur (D_L, L) défini par

$$D_L = \{\ u \in C^2[0,\pi] \mid u(0) = 0, u(\pi) + ku'(\pi) = 0\ \}$$
$$Lu = u'',\ u \in D_L$$

(a) Montrer que ses valeurs propres (λ_n) sont les solutions (quand elles existent) de l'équation $\operatorname{th}\sqrt{\lambda}\pi = -k\sqrt{\lambda}$

(b) Montrer que si λ_n est valeur propre, alors une fonction propre associée est de la forme $\phi_n(x) = A_n \operatorname{sh}\sqrt{\lambda_n}x$

(c) Montrer que $\lambda = 0$ n'est valeur propre que si $k = -\pi$ et que dans ce cas une fonction propre associée est de la forme $\phi_0(x) = A_0 x$.

7. On considère l'opérateur (D_L, L) défini par

$$D_L = \{\ u \in C^2[0,\pi] \mid u(0) = 0, u'(\pi) = 0\ \}$$
$$Lu = u'',\ u \in D_L$$

Déterminer les valeurs propres et les fonctions propres de (D_L, L).

8. Soit l'opérateur (D_L, L) défini par : $Lu = u''$ et D_L l'ensemble des fonctions u dans $C[0,\pi]$ vérifiant les conditions aux bords

$$\begin{cases} u'(0) + 2u'(\pi) = 0 \\ u(\pi) = 0 \end{cases}$$

(a) Montrer que (D_L, L) n'est pas symétrique et que si λ est réel, le problème : $\lambda u - Lu = 0, u \in D_L$, n'a aucune solution u non nulle.
(b) Montrer qu'il existe une infinité de valeurs complexes λ_n pour lesquelles le problème précédent possède une solution u_n non nulle. Déterminer les nombres (λ_n).
Solution : (a) Si l'opérateur (D_L, L) était symétrique, le wronskien de deux éléments quelconques de D_L serait nul, c'est-à-dire que l'on aurait $u(0)v'(\pi) = u'(\pi)v(0)$, quels que soient u et v dans D_L. La fonction définie par $u(x) = 1 + \cos x$ appartient à D_L et la condition précédente implique $v'(\pi) = 0$ et par suite $v'(0) = 0$ pour tout v dans D_L. Cela n'est pas vrai puisque, par exemple, la fonction définie par $v(x) = -\cos(x/2) + \sin(x)$ appartient bien à D_L, mais sa dérivée v' ne s'annule pas en π.
(b) La solution générale de l'équation $u'' - \lambda u = 0$ est de la forme $u(x) = A\,\mathrm{ch}(\sqrt{\lambda}x) + B\,\mathrm{sh}(\sqrt{\lambda}x)$. Donc une solution u appartient à D_L si et seulement si

$$A\,\mathrm{ch}(\sqrt{\lambda}\pi) + B\,\mathrm{sh}(\sqrt{\lambda}\pi) = 0 \quad \text{et} \quad B + 2A\,\mathrm{sh}(\sqrt{\lambda}\pi) + 2\,\mathrm{ch}(\sqrt{\lambda}\pi) = 0$$

c'est-à-dire si et seulement si $\mathrm{ch}(\sqrt{\lambda}\pi) = -2$. Il est clair que cette équation n'admet pas de solution réelle ; en posant $\sqrt{\lambda} = \mu + i\nu$, on vérifie rapidement que ses solutions sont de la forme

$$\sqrt{\lambda_n} = \mu + i(2n+1) \quad \text{avec} \quad n \in \mathbb{Z} \quad \text{et} \quad \mathrm{ch}\,\mu\pi = 2.$$

5.5 Etude spectrale des opérateurs de Sturm-Liouville

Le théorème 4.6 montre que si λ_0 n'est pas valeur propre de (D_L, L), le problème

$$\begin{cases} \lambda u - Lu = f \\ u \in D_L \end{cases}$$

est équivalent à l'équation intégrale

$$(\lambda - \lambda_0) G_{\lambda_0} u + u = G_{\lambda_0} f$$

qui fait intervenir l'opérateur intégral G_{λ_0} de noyau la fonction de Green $G_{\lambda_0}(.,.)$. Nous allons voir qu'il est possible de choisir λ_0 de façon que $G_{\lambda_0}(.,.)$ soit un noyau de type positif, on pourra alors appliquer les résultats du paragraphe 1 et notamment le théorème de Mercer. Les notations étant toujours celles du paragraphe 3, montrons d'abord le théorème suivant

Théorème 5.5.1. *L'opérateur (D_L, L) est semi-borné supérieurement. C'est-à-dire qu'il existe une constante M telle que, pour tout u dans D_L, on ait*

$$\langle Lu, u \rangle \leq M \|u\|^2$$

Démonstration. Par intégration par parties on obtient

$$\langle Lu, u \rangle = [pu'\overline{u}]_a^b - \int_a^b \left(p|u'|^2 + q|u|^2 \right) dx$$

Si les nombres θ et γ sont des multiples de $\pi/2$, le crochet est nul pour toute fonction u de D_L, donc pour une telle fonction

$$\langle Lu, u \rangle = - \int_a^b \left(p|u'|^2 + q|u|^2 \right) dx \leq M \|u\|^2$$

avec

$$M = -\inf\{\, q(x);\ a \leq x \leq b \,\}$$

Dans les autres cas, on a

$$\langle Lu, u \rangle = |u(b)|^2 \mathrm{cotg}\gamma - |u(a)|^2 \mathrm{cotg}\theta - \int_a^b \left(p|u'|^2 + q|u|^2 \right) dx$$

(en convenant de poser, pour $\theta = k\pi$, $|u(a)|^2 \mathrm{cotg}\theta = 0$). □

Lemme 5.5.2. *Soit u une fonction dans $C^2[a,b]$. Pour tout ϵ strictement compris entre 0 et $(a+b)/2$,*

$$\max_{a\leq x\leq b} |u(x)|^2 \leq \frac{2}{\epsilon}\int_a^b |u(y)|^2\,dy + 2\epsilon\int_a^b |u'(y)|^2\,dy$$

Démonstration. De l'égalité

$$u(x) - u(y) = \int_y^x u'(t)\,dt$$

on déduit, grâce à l'inégalité de Cauchy-Schwarz,

$$|u(x) - u(y)|^2 \leq |x-y| \int_a^b |u'(t)|^2\,dt$$

de plus

$$|u(x)|^2 \leq 2|u(x) - u(y)|^2 + 2|u(y)|^2$$

d'où par intégration par rapport à y sur $[x, x+\epsilon]$ ou $[x-\epsilon, x]$ (l'un au moins de ces deux intervalles est inclus dans $[a,b]$)

$$\epsilon|u(x)|^2 \leq 2\int_a^b |u(y)|^2\,dy + 2\epsilon^2\int_a^b |u'(t)|^2\,dt$$

Le lemme est ainsi prouvé. □

Terminons la preuve du théorème 5.1. Posons

$$A = \inf_{a\leq x\leq b} p(x), \quad B = \inf_{a\leq x\leq b} q(x) \quad \text{et} \quad C = \sup\{\ |\mathrm{cotg}\theta|, |\mathrm{cotg}\gamma|\ \}$$

Nous avons

$$\langle Lu, u\rangle \leq C\max|u|^2 - A\int_a^b |u'|^2\,dx - B\int_a^b |u|^2\,dx$$

D'après le lemme, pour tout $\epsilon > 0$

$$\langle Lu, u\rangle \leq (2C\epsilon - A)\int_a^b |u'|^2\,dx + \left(\frac{2C}{\epsilon} - B\right)\int_a^b |u|^2\,dx$$

Or, par hypothèse l'opérateur A est strictement positif ; on peut donc choisir ϵ de sorte que $2C\epsilon$ soit inférieur ou égal à A et il suffit alors de poser $M = 2C\epsilon^{-1} - B$. □

Corollaire 5.5.3. *Toute valeur propre de (D_L, L) est inférieure ou égale à $M - \inf\{ q(x); a \le x \le b \}$.*

Posons

$$m = \sup\{ \langle Lu, u\rangle \mid u \in D_L,\ \|u\| = 1 \}$$

D'après le théorème 5.1, m est fini, inférieur ou égal à M et toute valeur propre de (D_L, L) est inférieure ou égale à m.

Théorème 5.5.4. *Si $\lambda_0 > m$, la fonction de Green $G_{\lambda_0}(.,.)$ est un noyau de type positif.*

Démonstration. Soit f une fonction de $C[a,b]$ et soit $u = G_{\lambda_0} f$. D'après le théorème 4.5, on sait que $u \in D_L$ et $\lambda_0 u - Lu = f$ d'où

$$\langle G_{\lambda_0} f, f\rangle = \langle u, \lambda_0 u - Lu\rangle = \lambda_0 \|u\|^2 - \langle Lu, u\rangle \ge (\lambda_0 - m)\|u\|^2 \ge 0$$

ce qui est le résultat cherché. □

Théorème 5.5.5. (a) *Les valeurs propres de (D_L, L) constituent une suite (λ_n) qui tend vers $-\infty$.*

(b) *Les fonctions propres correspondantes, (ϕ_n), normalisées constituent une base hilbertienne de $L^2[a,b]$.*

(c) *Toute fonction u de D_L s'écrit*

$$u(x) = \sum_{n=1}^{\infty} \langle u, \phi_n\rangle \phi_n(x)$$

où la convergence est absolue et uniforme sur $[a,b]$.

(d) *Si λ n'est pas valeur propre de (D_L, L), le problème*

$$u \in D_L,\ \ \lambda u - Lu = f$$

admet, pour toute f continue, une solution unique.

(e) *Si λ est une valeur propre et ϕ une fonction propre correspondant à λ, le problème*

$$u \in D_L,\ \ \lambda u - Lu = f$$

admet une solution si, et seulement si, $\langle f, \phi\rangle = 0$.

(f) *Si λ n'est pas valeur propre, le noyau de la résolvante G_λ s'écrit*

$$G_\lambda(x,y) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(y)}{\lambda - \lambda_n}$$

la convergence étant absolue et uniforme sur $[a,b] \times [a,b]$.

Démonstration. Soit λ_0 un nombre réel qui n'est pas une valeur propre de (D_L, L). L'opérateur G_{λ_0} admet les mêmes sous-espaces propres que (D_L, L), ses valeurs propres sont les nombres

$$\mu = \frac{1}{\lambda_0 - \lambda}$$

où λ est une valeur propre de (D_L, L). D'après le théorème 2.9, chapitre IV, on sait que les valeurs propres de G_{λ_0} constituent une suite (μ_n), $n \geq 1$, infinie qui tend vers 0. On en déduit que les valeurs propres de (D_L, L) constituent une suite (λ_n) reliée à (μ_n) par la relation

$$\mu_n = \frac{1}{\lambda_0 - \lambda_n}$$

et qui tend donc vers $-\infty$ quand n tend vers l'infini. comme G_{λ_0} est un opérateur injectif, les fonctions propres (ϕ_n) correspondant à (λ_n) et normalisées, constituent une base hilbertienne de $L^2[a, b]$ par suite du théorème 3.4, chapitre IV.

La partie (c) de l'énoncé est une conséquence du théorème 1.3 de ce chapitre. La partie (d) a déjà été démontrée, c'est le (i) du théorème 4.6. De plus, le problème

$$u \in D_L, \ \lambda u - Lu = f$$

est, d'après ce même théorème, équivalent à l'équation intégrale

$$(\lambda - \lambda_0)G_{\lambda_0}u + u = G_{\lambda_0}f$$

Celle-ci admet une solution si, et seulement si, $G_{\lambda_0}f$ est orthogonale à ϕ (théorème 1.4). Comme

$$\langle G_{\lambda_0}f, \phi\rangle = \langle f, G_{\lambda_0}\phi\rangle = (\lambda - \lambda_0)\langle f, \phi\rangle$$

on en déduit la partie (e).

Pour $\lambda_0 > m$, le noyau $G_{\lambda_0}(x, y)$ est de type positif (théorème 5.3), donc d'après le théorème de Mercer (théorème 1.8)

$$G_{\lambda_0}(x, y) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(y)}{\lambda_0 - \lambda_n}$$

la convergence étant absolue et uniforme sur $[a, b] \times [a, b]$. Si λ n'est pas valeur propre, le rapport $(\lambda_0 - \lambda_n)/(\lambda - \lambda_n)$ est borné et par suite la série

$$G_{\lambda}(x, y) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(y)}{\lambda - \lambda_n}$$

converge également absolument et uniformément sur $[a, b] \times [a, b]$. Il est facile de montrer que sa somme est égale à $G_{\lambda}(x, y)$. Ce qui démontre la partie (f) de l'énoncé. □

EXERCICES

1. Soit l'opérateur (D_L, L), considéré à l'exercice 3 du paragraphe 4

$$D_L = \{ \; u \in C^2[0,\pi] \mid u(0) = 0, u(\pi) = 0 \; \}$$
$$Lu = u'', \;\; u \in D_L$$

(i) Ecrire le développement en série du noyau de Green de (D_L, L) suivant les fonctions propres (ϕ_n).
(ii) Que donne l'application du théorème de Mercer ?
(iii) Résoudre le problème

$$u \in D_L, \;\; (1/4)u + Lu = f$$

dans le cas où $f(x) = \sin 2x$ puis le cas où $f(x) = x/2$.

2. On considère l'opérateur

$$D_L = \{ \; u \in C^2[1,e] \mid u(1) = 0, u(e) = 0\}$$
$$Lu = x^2u'' + 2xu' + (1/4)u, u \in D_L$$

(a) Déterminer les valeurs propres (λ_n) et les fonctions propres (ϕ_n) de (D_L, L) (on peut effectuer le changement adéquat pour transformer L en un opérateur formellement auto-adjoint).
(b) Résoudre le problème

$$u \in D_L, \;\; Lu = x^{-\frac{1}{2}}$$

3. On reprend l'opérateur de l'exemple 4.9 du paragraphe précédent.
(a) Traduire le théorème de Mercer pour le noyau de Green G_λ.
(b) Appliquer la formule de la trace et en déduire l'égalité

$$\mathrm{cotg}\, z = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 + n^2\pi^2}$$

4. On considère l'opérateur (D_L, L) défini par

$$D_L = \{ \; u \in C^2[0,1] \mid u(0) = 0, u(1) + u'(1) = 0 \; \}$$
$$Lu = u'' - u, \;\; u \in D_L$$

(a) Montrer que (D_L, L) est injectif et déterminer son noyau de Green.
(b) Trouver la relation qui lie les valeurs propres (λ_n) de (D_L, L) aux valeurs propres (μ_n) de l'opérateur de Green G.
(c) Montrer que les fonctions propres de (D_L, L) forment une base hilbertienne de $L^2[0,1]$.
(d) Déterminer la somme de la série $\sum \lambda_n^{-2}$.

5. Reprendre les questions de l'exercice 5 pour les deux cas suivants

$$D_L = \{ \ u \in C^2[0,1] \mid u(0) = u'(0), u(1) + u'(1) = 0 \ \}$$
$$Lu = u'' - u, \ \ u \in D_L \qquad (a)$$

$$D_L = \{ \ u \in C^2[0,1] \mid u'(0) = 0, u'(1) = 0 \ \}$$
$$Lu = u'', \ \ u \in D_L \qquad (b)$$

5.6 Etude spectrale de l'opérateur de Bessel

Dans ce paragraphe nous présentons un exemple d'opérateur de Sturm-Liouville singulier. Il s'agit de l'opérateur de Bessel dont l'étude nous est accessible grâce au fait que sa résolvante est un opérateur de Hilbert-Schmidt. On pourrait mener une étude similaire pour des opérateurs de Sturm-Liouville singuliers, dont la résolvante est compacte. C'est le cas, par exemple, de l'opérateur de Legendre et l'opérateur d'Hermite.

Soit $L^2((0,1), xdx)$ l'espace de Hilbert des (classes de) fonctions définies et de carré intégrables sur l'intervalle ouvert $(0,1)$, relativement à la mesure xdx. Le produit scalaire et la norme y sont définis par

$$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\, xdx \quad \text{et} \quad \|f\|^2 = \int_0^1 |f(x)|^2\, xdx$$

Dans l'espace $L^2((0,1), xdx)$, on considère l'opérateur intégral G dont le noyau k est défini par

$$k(x,y) = \begin{cases} \operatorname{Log} x, & \text{si } 0 < y \leq x < 1\,; \\ \operatorname{Log} y, & \text{si } 0 < x \leq y < 1. \end{cases}$$

Le noyau k a été évoqué dans l'exercice 3 du paragraphe 1. On peut vérifier que k est de carré intégrable sur $(0,1) \times (0,1)$, relativement à la mesure produit $xydxdy$. Mieux encore,

$$\int_0^1 |k(x,y)|^2\, ydy = \frac{x^2}{2} |\operatorname{Log} x|^2 + \int_x^1 |\operatorname{Log} y|^2\, ydy$$

Comme la fonction $y \mapsto y|\operatorname{Log} y|^2$ est bornée sur $(0,1)$, on en déduit qu'il existe $M > 0$, tel que

$$\sup_{0<x<1} \int_0^1 |k(x,y)|^2\, ydy < M^2 \qquad (5.7)$$

L'opérateur G de noyau k est donc un opérateur de Hilbert-Schmidt ; il est auto-adjoint car k est symétrique.

Proposition 5.6.1. *Pour toute f dans $L^2((0,1), xdx)$,*

(a) *la fonction Gf est continue et bornée sur $(0,1)$; elle vérifie deplus :* $\lim_{x \to 1^-} Gf(x) = 0$.

(b) *Si de plus f est continue alors $u = Gf$ est deux fois dérivable sur $(0,1)$ et*

$$\frac{1}{x}\frac{d}{dx}\left(x\frac{du}{dx}\right) = f \quad \text{et} \quad \lim_{x \to 0^+} xu'(x) = 0.$$

Démonstration. Si l'on pose $u = Gf$, il vient

$$u(x) = \operatorname{Log} x \int_0^x f(y)y\,dy + \int_x^1 \operatorname{Log} y f(y)y\,dy$$

et si f est continue

$$u'(x) = \frac{1}{x}\int_0^x f(y)y\,dy \quad \text{et} \quad (xu')'(x) = xf(x)$$

La proposition s'en déduit immédiatement. □

Nous allons voir que G est "la résolvante" d'un opérateur de Sturm-Liouville "singulier", appelé opérateur de Bessel.
Soit D_L l'ensemble des fonctions u dans $C^2(0,1)$ qui sont bornées et telles que $\lim_{x\to 1} u(x) = 0$ et $\lim_{x \to 0} xu'(x) = 0$. Pour $u \in D_L$, on pose $Lu = x^{-1}(xu')'$. (D_L, L) est un opérateur de Sturm-Liouville singulier, puisque le coefficient de u' s'annule en 0 ; il est formellement auto-adjoint dans l'espace $L^2((0,1), xdx)$.

Conformément aux notations adoptées dans ce chapitre (voir le début du paragraphe 3), on pose $[u,v] = x(uv' - vu')$. On vérifie rapidement que les fonctions $u_1(x) = 1$ et $u_2(x) = \operatorname{Log} x$ sont deux solutions de $Lu = 0$ et $[u_1, u_2] = 1$. On comprend maintenant l'origine du noyau k : il s'exprime à l'aide de u_1 et u_2 par la même formule (1) du paragraphe 4.

Proposition 5.6.2. *L'opérateur (D_L, L) est injectif.*

Démonstration. En effet, toute solution u de $Lu = 0$, non identiquement nulle, est une combinaison linéaire de u_1 et u_2 et ne peut donc satisfaire les conditions aux limites. □

Proposition 5.6.3. *Avec les notations de ce paragraphe,*

(i) *Si f est continue sur $(0,1)$, alors $u = Gf$ appartient à D_L et satisfait l'égalité $Lu = f$.*

(ii) *Inversement, soit u dans D_L et posons $f = Lu$, alors f est continue sur $(0,1)$ et $Gf = u$.*

(iii) *Pour u et v dans D_L, on a $\langle Lu, v\rangle = \langle u, Lv\rangle$.*

Démonstration. L'assertion (i) est une reformulation de la proposition 6.1 (b). Soit $u \in D_L$, la fonction $f = Lu$ est évidemment continue sur $(0, 1)$. D'après l'assertion (b) de la proposition 6.1, la fonction $v = Gf$ est dans D_L et vérifie $Lv = f$. Comme (D_L, L) est injectif on en déduit que $v = u$. Pour montrer l'assertion (iii), on remarque d'abord que, si u est dans D_L, alors xu' est bornée sur $(0, 1)$, car une telle fonction s'écrit (d'après (ii)) sous la forme $u = Gf$, avec $f = Lu$ et par suite

$$xu'(x) = \int_0^x f(y)\, ydy$$

Maintenant, pour u et v dans D_L, on a

$$\langle Lu, v\rangle = \int_0^1 (xu')'\overline{v}\, dx = [u, \overline{v}](1) - [u, \overline{v}](0) + \langle u, Lv\rangle$$

Le fait que u et v soient dans D_L et ce qui précède montrent que les deux premiers termes du dernier membre sont nuls. □

La proposition 6.3 exprime le fait que l'opérateur G est une bijection de $C(0, 1)$ sur D_L et a pour "inverse" l'opérateur L. Autrement dit, G est la résolvante de l'opérateur de Sturm-Liouville (D_L, L).

Proposition 5.6.4. *L'opérateur G est injectif de l'espace $L^2((0, 1), xdx)$ dans lui-même.*

Démonstration. Compte tenu de la relation $(\Im m G)^\perp = \ker(G^*)$ et de ce que G est auto-adjoint, on doit montrer que l'image de G est dense dans $L^2((0, 1), xdx)$. Or, la proposition 6.3 montre que l'image de G contient D_L. Comme ce dernier est dense dans $L^2((0, 1), xdx)$ (noter que l'espace des fonctions C^∞ à support compact dans $(0, 1)$ contient D_L), on en déduit le résultat voulu. □

Soit ϕ une fonction propre de G ; cela veut dire que ϕ est non identiquement nulle, appartient à $L^2((0, 1), xdx)$ et il existe une constante μ (non nulle car G est injectif), telle que $G\phi = \mu\phi$. Compte tenu des propositions 6.1 et 6.3, on déduit de cette égalité que ϕ est dans D_L et que $\phi = \mu L\phi$, c'est-à-dire que ϕ est fonction propre de (D_L, L) associée à la valeur propre μ^{-1}. Inversement, on montre que si ϕ est une fonction propre de (D_L, L) associée à la valeur propre λ ($\lambda \neq 0$ d'après la proposition 6.2), alors ϕ est fonction propre de G associée à la valeur propre λ^{-1}.

• L'opérateur G étant injectif, ses fonctions propres, qui sont aussi les fonctions propres de l'opérateur (D_L, L), forment une base hilbertienne de l'espace $L^2((0, 1), xdx)$. Nous allons déterminer cette base hilbertienne et les valeurs propres correspondantes. Soit ϕ une fonction propre associée à une valeur propre μ ($\mu \neq 0$ d'après ce qui précède). En multipliant les

deux membres de $L\phi = \mu^{-1}\phi$ par $\overline{\phi}$ et en intégrant sur $(0,1)$, il vient $\langle L\phi, \overline{\phi}\rangle = \mu^{-1}\|\phi\|^2$. Intégrant par parties et tenant compte de (f), il vient

$$-\int_0^1 |\phi'(x)|^2\, x dx = \mu^{-1}\|\phi\|^2$$

Il en résulte que toute valeur propre μ est négative. Posons $\mu = -\lambda^{-2}$ et faisons le changement de variable $t = \lambda x$. On vérifie rapidement que la fonction ψ, définie par $\psi(t) = \phi(x)$, est solution de

$$t\psi'' + \psi' + t\psi = 0 \tag{5.8}$$

C'est l'équation différentielle de Bessel d'indice 0. Cherchons une solution de cette équation sous la forme d'une série entière

$$\psi(t) = \sum_{k=0}^{\infty} a_k t^k$$

où les coefficients a_k sont à déterminer de façon que ϕ soit dans D_L. On a

$$\begin{aligned}
t\psi &= \qquad a_0 t + a_1 t^2 + a_2 t^3 + \cdots + a_{n-1}t^n + \cdots \\
\psi' &= a_1 + 2a_2 t + 3a_3 t^2 + 4a_4 t^3 + \cdots + (n+1)a_{n+1}t^n + \cdots \\
t\psi'' &= \qquad 2a_2 t + 3.2a_3 t^2 + 4.3a_4 t^3 + \cdots + (n+1)na_{n+1}t^n + \cdots
\end{aligned}$$

En ajoutant membre à membre ces trois égalités et en remarquant l'identité $(n+1)n + (n+1) = (n+1)^2$, on obtient

$$t\psi'' + \psi' + t\psi = a_1 + (2^2 a_2 + a_0)t + \cdots + ((n+1)^2 a_{n+1} + a_{n-1})t^n + \cdots$$

On en déduit que $a_{2k+1} = 0$, pour tout entier k, et les coefficients d'indice pair vérifient la relation de récurrence

$$(2k)^2 a_{2k} + a_{2k-2} = 0, \quad \text{pour} \quad k \in \mathbb{N}$$

En prenant $a_0 = 1$, on trouve la solution

$$J_0(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k!)^2}\left(\frac{t}{2}\right)^{2k}$$

c'est la fonction de Bessel d'indice 0. Le rayon de convergence de la série du second membre est infini, et par suite la fonction de Bessel J_0 est une solution entière de l'équation (2). Elle vérifie de plus $J_0(0) = 1$ et $J_0'(0) = 0$. Montrons que les seules solutions bornées sur l'intervalle $(0,1)$ sont celles qui sont proportionnelles à J_0. En effet, la formule (2) du paragraphe 2

montre qu'une deuxième solution linéairement indépendante de J_0 est donnée par

$$u(t) = J_0(t)\int \frac{ds}{sJ_0^2(s)}$$

La fonction J_0^{-2} est analytique au voisinage de 0 et donc on peut écrire dans un tel voisinage

$$\frac{1}{tJ_0^2(t)} = \frac{1}{t} + \text{une série entière en } t \text{ convergente au voisinage de } 0$$

On en déduit que la deuxième solution u peut s'écrire sous la forme

$$u(t) = J_0(t)\,\mathrm{Log}\,(t/2) + v(t)$$

où la fonction v doit vérifier l'équation différentielle

$$v'' + \frac{1}{t}v' + v = \frac{2}{t}J_0'$$

Si l'on pose

$$v(t) = \sum_{k=1}^{\infty}(-1)^k\frac{b_k}{(k!)^2}\left(\frac{t}{2}\right)^{2k}$$

les coefficients (b_k) doivent vérifier la relation de récurrence

$$b_{k+1} = b_k + 1/(k+1)$$

En prenant $b_1 = 1$, on obtient $b_k = 1 + 2^{-1} + \cdots + k^{-1}$ et le rayon de convergence de la série définissant v est infini. Ainsi, toute solution, sur $(0, +\infty)$, de l'équation (2) est de la forme

$$u(t) = AJ_0(t) + B\Big[J_0(t)\,\mathrm{Log}\Big(\frac{t}{2}\Big) + v(t)\Big]$$

où A et B sont des constantes. On appelle fonction de Bessel de deuxième espèce et d'ordre 0, et on note Y_0, la solution définie par

$$Y_0(t) = \frac{2}{\pi}\Big(\mathrm{Log}(t/2) + \gamma\Big)J_0(t) + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{(n!)^2}\Big(1 + \frac{1}{2} + \cdots + \frac{1}{n}\Big)(t/2)^{2n}$$

où γ est la constante d'Euler[9],

$$\gamma = \Big(\lim_{p\to\infty}\sum_{1}^{p} n^{-1} - \mathrm{Log}\,p\Big) = 0.577215...$$

[9]Leonhard Euler (1707–1783) est né à Bâle, en Suisse. Avec Joseph-Louis Lagrange, son émule plus jeune, Euler est l'un des deux géants mathématiques qui ont dominé la science du XVIIIe siècle. Ses travaux, d'une abondance inégalée, couvrent tout champ

En revenant à la variable x, on peut dire que toute solution sur $(0,1)$ de l'équation $Lu = \lambda u$ est de la forme

$$u(x) = AJ_0(\lambda x) + B\Big[J_0(\lambda x)\operatorname{Log}\Big(\frac{\lambda}{2}\Big) + v(\lambda x)\Big]$$

Si ϕ est une fonction propre de (D_L, L), ϕ est bornée et nulle en 1, donc nécessairement $B = 0$ et $J_0(\lambda) = 0$. On en déduit que les valeurs propres de l'opérateur (D_L, L), qui sont aussi les inverses des valeurs propres de l'opérateur G, sont de la forme $-\lambda_n^2$, où (λ_n) est la suite des zéros de la fonction $\lambda \mapsto J_0(\lambda)$.

Le graphe de la fonction $x \mapsto J_0(x)$, représenté ci-dessous, montre que la fonction de Bessel J_0 est toujours majorée par 1 et qu'elle atteint cette valeure en $x = \pm 1$. La preuve de cette propriété remarquable est donnée dans l'exercice 1. Le comportement de $J_0(x)$, lorsque x tend vers l'infini, est donné par

$$J_0(x) \simeq \sqrt{2/(\pi x)}\cos(x - \frac{\pi}{4})$$

si bien que la courbe d'équation $y = \sqrt{2/(\pi x)}$ représente l'enveloppe de celle de J_0.

Le théorème suivant est l'analogue des théorèmes 4.3 et 5.3 du chapitre II, qui concernent les opérateurs de Legendre et d'Hermite.

Théorème 5.6.5. *Les valeurs propres de (D_L, L) sont $(-\lambda_n^2)$, où les (λ_n) sont les zéros positifs de la fonction J_0. Les sous-espaces propres correspondants sont de dimension 1. Les fonctions propres correspondantes forment une base hilbertienne de l'espace $L^2((0,1), xdx)$ et sont données par :*

$$\phi_n(x) = \sqrt{2}\frac{J_0(\lambda_n x)}{J_0'(\lambda_n)}$$

Démonstration. L'opérateur de Hilbert-Schmidt G étant injectif, ses fonctions propres (normalisées), qui sont aussi les fonctions propres de (D_L, L), forment une base hilbertienne de l'espace $L^2((0,1), xdx)$. Ce qui précède

des mathématiques, de la mécanique céleste et de la physique de son époque. Il a renouvelé l'articulation entre les secteurs mathématiques, fixé la plupart des notations du calcul infinitésimal que nous utilisons encore, développé la théorie des nombres de Fermat et systématisé la géométrie analytique de Descartes tout en l'étendant du plan à l'espace. Lors de l'étude des développements asymptotiques de sommes partielles de séries divergentes, il a découvert le nombre γ, qui porte son nom, et a donné une méthode pour trouver une valeur approchée de γ avec 15 décimales exactes. Mais ce nombre reste encore de nature mystérieuse et on ne sait pas encore s'il est rationnel ou irrationnel.

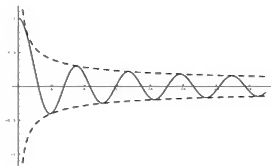


FIG. 5.1 – La fonction de Bessel J_0 et son enveloppe

montre que ces fonctions propres ϕ_n sont de la forme $\phi_n(x) = c_n J_0(\lambda_n x)$. Les coefficients c_n sont choisis de façon que ϕ_n soit de norme 1, c'est-à-dire $\int_0^1 |\phi_n(x)|^2 \, x dx = 1$. Pour calculer c_n, soient λ et ν dans $\mathbb{C}$, les fonctions $u(x) = J_0(\lambda x)$ et $v(x) = J_0(\nu x)$ vérifient

$$\langle Lu, v\rangle - \langle u, Lv\rangle = [u, \overline{v}](1)$$

c'est-à-dire

$$-(\lambda^2 - \nu^2)\langle u, v\rangle = \lambda J_0'(\lambda) J_0(\nu) - \nu J_0'(\nu) J_0(\lambda)$$

Supposons que λ soit un zéro de J_0. En divisant les deux membres de la deuxième égalité par $(\lambda^2 - \nu^2)$ et en faisant tendre ν vers λ, nous obtenons

$$\int_0^1 [J_0(\lambda x)]^2 \, x dx = \frac{1}{2}[J_0'(\lambda)]^2$$

L'expression de ϕ_n s'en déduit immédiatement. □

Corollaire 5.6.6. *Toute fonction f de $L^2((0,1), x dx)$ se développe en série suivant les fonctions de Bessel :*

$$f(x) = \sum_{n=0}^{\infty} c_n(f) J_0(\lambda_n x), \quad c_n(f) = \frac{2}{[J_0'(\lambda_n)]^2} \int_0^1 f(t) J_0(\lambda_n t) \, t dt$$

où la série converge vers f dans $L^2((0,1), xdx)$. De plus, si f et g sont dans $L^2((0,1), xdx)$, on a la formule de Parseval

$$\int_0^1 f(x)\overline{g(x)}\,xdx = \sum_{n=0}^{\infty} \frac{[J_0'(\lambda_n)]^2}{2} c_n(f)\overline{c_n(g)}.$$

Le développement de f suivant les fonctions de Bessel est parfois appelé le développement de Fourier-Bessel de f, et les coefficients $c_n(f)$ sont appelés les coefficients de Fourier-Bessel de f.

Exemple 5.6.7. - Soient $\mathbf{1}$ la fonction constante égale à 1 et $u(x) = x^2$. Nous allons chercher les développements de Fourier-Bessel de ces deux fonctions. En remarquant que $(x\phi_n')' = -\lambda_n^2 x\phi_n$, on a

$$\langle \mathbf{1}, \phi_n\rangle = \int_0^1 x\phi_n(x)\,dx = -\lambda_n^{-2}\int_0^1 (x\phi_n')'\,dx$$
$$\langle u, \phi_n\rangle = \int_0^1 x^3\phi_n(x)\,dx = -\lambda_n^{-2}\int_0^1 x^2(x\phi_n')'\,dx$$

En tenant compte du fait que $\phi_n(1) = 0$ et que $\phi_n'(x) = \lambda_n J_0'(\lambda_n x)$, des intégrations par parties donnent

$$\langle \mathbf{1}, \phi_n\rangle = -\frac{J_0'(\lambda_n)}{\lambda_n^2} \quad \text{et} \quad \langle u, \phi_n\rangle = \frac{4-\lambda_n^2}{\lambda_n^3} J_0'(\lambda_n)$$

On en déduit les développements en séries de Fourier-Bessel suivants

$$1 = \sum_{n=0}^{\infty} \frac{-2}{\lambda_n J_0'(\lambda_n)} J_0(\lambda_n x), \quad x^2 = \sum_{n=0}^{\infty} \frac{2(4-\lambda_n^2)}{\lambda_n^3 J_0'(\lambda_n)} J_0(\lambda_n x)$$

où la convergence a lieu dans l'espace $L^2((0,1), xdx)$. On peut en déduire le développement suivant

$$1 - x^2 = \sum_{n=0}^{\infty} \frac{-8}{\lambda_n^3 J_0'(\lambda_n)} J_0(\lambda_n x)$$

La formule de Parseval appliquée, d'une part à $f = g = \mathbf{1}$, et d'autre part à $f = \mathbf{1}$ et $g = 1 - x^2$ donne

$$\frac{1}{4} = \sum_{n=0}^{\infty} \frac{1}{\lambda_n^2} \quad \text{et} \quad \int_0^1 (1-x^2)\,xdx = \sum_{n=0}^{\infty} \frac{8}{\lambda_n^4}$$

Sachant que les valeurs propres de l'opérateur intégral G sont les $(-\lambda_n^{-2})$, la dernière relation implique la suivante

$$|||G|||^2 = \sum_{n=0}^{\infty} \frac{1}{\lambda_n^4} = \frac{1}{32}$$

Notons que le premier membre de légalité ci-dessus, peut se calculer directement en utilisant l'expression du noyau de l'opérateur G.

EXERCICES

1. On pose $I_n = \int_0^\pi \sin^{2n}\theta\, d\theta$.
(a) Montrer que pour $n \geq 1$, $I_n = ((2n-1)/2n))I_{n-1}$. Sachant que $I_0 = \pi$, en déduire par récurrence que

$$I_n = \frac{(2n)!}{2^{2n}(n!)^2}\pi$$

(b) En utilisant ce qui précède, le développement en série entière de $\cos u$ et celui de J_0 donné par (2), montrer que

$$J_0(x) = \frac{1}{\pi}\int_0^\pi \cos(x\sin\theta)\, d\theta$$

(c) En déduire la représentation intégrale de $J_0(\lambda x)$, pour $\lambda \in \mathbb{C}$ et montrer que, pour tout λ dans $\mathbb{R}$,

$$|J_0(\lambda x)| \leq 1$$

2. Soient a un réel positif et $E = L^2((0,a), xdx)$ l'espace de Hilbert des (classes de) fonctions de carré intégrables pour la mesure xdx. On désigne par D_L le sous-espace des fonctions $u \in C^2(0,a)$ qui sont bornées sur $(0,a)$ et telle que : $\lim_{x\to a} u(x) = 0$ et $\lim_{x\to 0} xu'(x) = 0$. Pour u dans D_L, on pose : $Lu = (1/x)\big(xu'\big)'$.
(a) Déterminer une base de solutions de l'équation $Lu = 0$ et en déduire que l'opérateur (D_L, L) est injectif.
(b) Déterminer le noyau de la résolvante de (D_L, L).
(c) Déterminer une base hilbertienne de $L^2(0,a), xdx)$, formée de fonctions propres de l'opérteur (D_L, L). (On peut adopter la méthode suivie dans le cas de l'intervalle $(0,1)$; il est peut être plus intéressant de faire un changement de variables qui permet de se ramener au cas de l'intervalle $(0,1)$).

ANNEXE

Afin de rendre plus facile la lecture de ce livre, nous avons regroupé dans cette annexe les principaux théorèmes utilisés, avec des démonstrations plus ou moins complètes.

1. Espace de Banach

1.1 DÉFINITION. *Un espace de Banach est un espace vectoriel normé complet pour la métrique définie à partir de sa norme.*

Le critère suivant, bien que facile à démontrer, est très utile

1.2 PROPOSITION. *Soit E un espace vectoriel normé. Les trois propriétés suivantes de E sont équivalentes*

(a) *E est un espace de Banach*
(b) *Toute série d'éléments de E normalement convergente est convergente*
(c) *Toute série $\sum x_n$, $x_n \in E$, telle que $\|x_n\| \leq 2^{-n}$ pour tout $\in \mathbb{N}$ est convergente.*

Rappelons qu'une série est normalement convergente si la série à termes positifs $\sum_{n \geq 0} \|x_n\|$ converge.

Démonstration. — Le fait que (a) implique (b) est clair car, les sommes partielles $S_n = x_1 + x_2 \cdots + x_n$ forment alors une suite de Cauchy qui doit converger. De même, (b) implique de façon évidente (c). Pour montrer que (c) implique (a) soit (y_n) une suite de Cauchy. Il existe une sous-suite $z_k = y_{n(k)}$ telle que $\|z_k - z_{k+1}\| \leq 2^{-k}$. On pose alors

$$x_0 = z_0, x_1 = z_1 - z_0, \ldots, x_k = z_k - z_{k+1}, \ldots$$

l'assertion (c) entraîne l'existence d'un élément x de E, tel que

$$\lim_{k \to \infty} \|x - z_k\| = 0$$

Pour tout $\epsilon > 0$, il existe un entier k_0 tel que, si $k > k_0$, on ait $\|x - y_{n(k)}\| \leq \epsilon/2$. Puisque (y_n) est une suite de Cauchy, il existe n_0

tel que, si $n > n_0$ et $m > n_0$, on ait $\|y_n - y_m\| \leq \epsilon/2$. En combinant ces deux inégalités quand $n(k) > n_0$, il vient $\|x - y_m\| \leq \epsilon$ pour $m \geq n_0$. □

1.3 Remarque.

La proposition dit que si un espace vectoriel normé n'est pas complet, on n'y dispose d'aucun critère de convergence pour les séries, c'est là un avantage de travailler dans un espace de Banach. Cependant, un espace de Banach contient toujours des éléments *difficiles à appréhender*, car il contient tous les éléments que l'on peut fabriquer à l'aide de séries normalement convergentes.

1.4 Exemple.

On désigne par $C[0,1]$ l'espace vectoriel des fonctions continues sur $[0,1]$, à valeurs dans $\mathbb{C}$. Pour f dans $C[0,1]$, on pose

$$\|f\|_\infty = \sup_{0\leq x\leq 1} |f(x)|$$

on sait que ce maximum est fini et atteint. Alors $C[0,1]$, muni de cette norme est un espace de Banach.

Démonstration. — Soit $\sum_{n\geq 1} f_n$ une série normalement convergente de fonctions continues sur $[0,1]$. Puisque $|f_n(x)| \leq \|f_n\|_\infty$, la série numérique de terme général $f_n(x)$ est absolument convergente, soit $S(x)$ sa somme. On définit ainsi une fonction S sur $[0,1]$ et on a deux choses à prouver : la continuité de S et la convergence (au sens de la norme) de la série $\sum_{n\geq 0} f_n$ vers S. Soit n un entier fixé assez grand pour que l'on ait

$$\sum_{k=n+1}^{\infty} \|f_k\|_\infty \leq \epsilon$$

et soit $S_n = f_1 + f_2 + \cdots + f_n$. Alors, S_n est continue sur $[0,1]$ et il existe $\eta > 0$ tel que $|x - y| \leq \eta$ implique $|S_n(x) - S_n(y)| \leq \epsilon$. Finalement, l'inégalité $|x - y| \leq \eta$ implique

$$|S(x) - S(y)| \leq |S(x) - S_n(x)| + |S_n(x) - S_n(y)| + |S_n(y) - S(y)| \leq 3\epsilon$$

Le fait que $\|S - S_n\|_\infty$ tende vers 0, lorsque n tend vers l'infini, est clair car on a, pour tout x dans $[0,1]$,

$$|S(x)-S_n(x)| \leq \sum_{k=n+1}^{\infty} |f_k(x)|, \quad \text{et donc} \quad \|S-S_n\|_\infty \leq \sum_{k=n+1}^{\infty} \|f_k\|_\infty. \quad \square$$

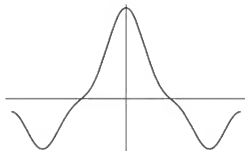


Figure A.

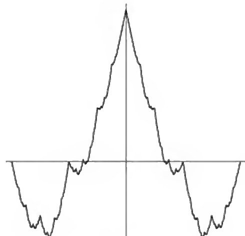


Figure B.

C'est parce que $C[0,1]$ est complet qu'il contient des fonctions qui, bien que continues, sont tout de même *difficiles à saisir*. Par exemple la fonction de Weierstrass définie par

$$f(x) = \sum_{n=0}^{\infty} 2^{-n} \cos(a^n x), \quad a \geq 2,\ a \in \mathbb{N}$$

n'est nulle part dérivable. Elle n'est croissante sur aucun intervalle (si petit soit-il), et elle n'est décroissante sur aucun intervalle (si petit soit-il). La figure A (resp. B) précédente représente le graphe de la somme

partielle d'ordre 2 (resp. d'ordre 10) de la série définissant la fonction de Weierstrass avec $a = 2$.

De même, l'espace vectoriel $C^m[0,1]$ des fonctions dont toutes les dérivées jusqu'à l'ordre m sont continues sur $[0,1]$, muni de la norme

$$\|f\| = \sup\{ \|f^{(r)}\|_\infty, 0 \leq r \leq m \}$$

est un espace de Banach.

1.5 Exemple.

Soit (X, Ω, μ) un espace mesuré σ-fini. Si $1 \leq p < \infty$, on désigne par $L^p(X, \Omega, \mu)$ l'ensemble des (classes de) fonctions f telles que

$$\|f\|_p = \left(\int_X |f(x)|^p \, d\mu\right)^{1/p} < \infty$$

et on désigne par $L^\infty(X, \Omega, \mu)$ l'ensemble des (classes de) fonctions f essentiellement bornées, c'est-à-dire telles que

$$\|f\|_\infty = \operatorname{supess} |f(x)| < \infty$$

Si p est strictement compris entre 1 et l'infini, on définit son exposant conjugué q par la relation $(1/p) + (1/q) = 1$ et on convient que $p = 1$ et $+\infty$ sont conjugués. L'application $f \mapsto \|f\|_p$ vérifie, pour tout $1 \leq p \leq \infty$,
L'inégalité de Hölder : pour $f \in L^p$ et $g \in L^q$, $\|fg\|_1 \leq \|f\|_p \|g\|_q$.
L'inégalité de Minkowski : pour f et g dans L^p, $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.
On en déduit alors que $f \mapsto \|f\|_p$ est une norme sur $L^p(X, \Omega, \mu)$. En utilisant la proposition 2, on montre que $L^p(X, \Omega, \mu)$ muni de cette norme est un espace de Banach.

1.6 Remarque.

On rappelle que, si μ est une mesure de Radon sur $\mathbb{R}^n$, l'espace $C_c(\mathbb{R}^n)$ des fonctions continues à support compact est dense dans $L^p(\mathbb{R}^n, d\mu)$ pour $1 \leq p < \infty$. Si I est un intervalle borné, $C(\overline{I})$ est dense dans $L^p(I, dx)$. Cela permet d'ailleurs de définir abstraitement les espaces L^p sans théorie de l'intégration préalable. On part, en effet, de l'espace vectoriel $C(\overline{I})$ (respectivement $C_c(\mathbb{R}^n)$), qu'on munit de la norme $\|f\|_p$ (c'est bien une norme car une fonction continue positive d'intégrale nulle est identiquement nulle). On construit alors $L^p(I, dx)$ (respectivement

$L^p(\mathbb{R}^n, dx)$) par le procédé canonique de complétion (celui qui permet de construire $\mathbb{R}$ à l'aide de l'ensemble de toutes les suites de Cauchy de nombres rationnels).

2. Prolongement des applications linéaires continues

Soient E et F deux espaces métriques et $E' \subset E$ un sous-ensemble dense de E. On s'intéresse à l'étude des propriétés d'une application f de E dans F, connaissant les propriétés de sa restriction $f|_{E'}$ de E' dans F. Cela joue un rôle fondamental, en effet de même qu'une calculatrice ne manie que des nombres rationnels (et jamais des nombres réels), en analyse on ne manipule jamais les fonctions "générales" des espaces de Banach (tels que $L^p(\mathbb{R})$). On ne traite que les approximations (telles que $C_c(\mathbb{R})$ ou mieux encore $C_c^\infty(\mathbb{R})$).

2.1 Théorème. *Soient E un espace vectoriel normé et E' une partie dense dans E. Soient F un espace vectoriel normé complet, C une constante et $A : E' \to F$ une application linéaire telle que*

$$\|Ax\| \leq C\|x\|, \quad \forall x \in E'$$

Alors, il existe une et une seule application linéaire continue $\tilde{A} : E \to F$, dont la restriction à E' coïncide avec A, de plus

$$\|\tilde{A}x\| \leq C\|x\|, \quad \forall x \in E$$

Démonstration. — La preuve est trés simple. Soit x dans E, il existe au moins une suite (x_n) dans E' telle que $\|x - x_n\| \to 0, (n \to \infty)$. Alors (x_n) est une suite de Cauchy et il en est de même de $(y_n = Ax_n)$. Puisque F est complet, (y_n) converge vers un élément y quand n tend vers l'infini. Pour poser $y = \tilde{A}x$, il suffit de s'assurer de ce que y ne dépende pas de la suite (x_n) choisie. Mais si (x'_n) en est une autre, $\|x_n - x'_n\|$ tend vers 0 quand n tend vers l'infini et il en sera de même de $\|y_n - y'_n\|$. Enfin, on vérifie rapidement que $\tilde{A}$ est linéaire et, par passage à la limite dans les inégalités, que $\|\tilde{A}x\| \leq C\|x\|$. □

2.2 Remarque.

Si E' est dense dans E et si A est une application continue de E' dans F, il n'existe pas nécessairement d'application continue de E dans F qui prolonge A. Cependant, si on suppose que A est uniformément

continue (propriété automatiquement satisfaite si A est linéairte), alors le prolongement existe et est unique. La linéarité joue donc un rôle important, notamment dans les théorèmes de Banach que nous allons présenter dans la section qui suit.

3. Les théorèmes de Banach

Une propriété agréable des espaces complets est que la théorie des opérateurs linéaires y est plus simple. La manière dont on exploite souvent le fait que l'espace est complet repose sur le théorème suivant, valable sur les espaces métriques complets.

3.1 Théorème de Baire. *Soit E un espace métrique complet. Alors l'une des deux propriétés équivalentes suivantes a lieu*

(a) *Pour toute famille dénombrable d'ouverts U_n dense dans E, $\bigcap_n U_n$ est dense dans E.*

(b) *Pour toute famille dénombrable de fermés F_n d'intérieur vide, la réunion $\bigcup_n F_n$ a un intérieur vide.*

La version suivante du théorème de Baire est souvent utilisée

3.2 Théorème. *Soit E un espace métrique complet. Supposons que E soit égal à une réunion dénombrable d'ensembles fermés $E = \bigcup_n F_n$. Alors un au moins de ces fermés est d'intérieur non vide.*

Démonstration. — Il est clair que les propriétés (a) et (b) sont équivalentes par passage au complémentaire. Montrons (a), c'est-à-dire que $U = \bigcap_n U_n$ est dense dans E. Cela revient à prouver que, pour tout $x \in E$ et tout $\epsilon > 0$,

$$U \cap B(x, \epsilon) \neq \emptyset$$

où $B(x, \epsilon)$ désigne la boule ouverte de centre x et de rayon ϵ. Puisque U_1 est dense dans E, l'intersection $B(x, \epsilon) \cap U_1$ est un ouvert non vide et il existe une boule $B(x_1, \epsilon_1)$ telle que

$$\overline{B}(x_1, \epsilon_1) \subset B(x, \epsilon) \cap U_1, \quad \text{avec} \quad \epsilon_1 < 2^{-1}$$

On construit la boule $B(x_2, \epsilon_2)$ de la façon suivante : puisque U_2 est dense dans E, l'ouvert $B(x_1, \epsilon_1) \cap U_2$ est non vide et contient donc une boule $B(x_2, \epsilon_2)$ telle que

$$\overline{B}(x_2, \epsilon_2) \subset B(x_1, \epsilon_1) \cap U_2, \quad \text{avec} \quad \epsilon_2 < 2^{-2}$$

Par récurrence, ce procédé permet de construire, pour tout n, la boule $B(x_n, \epsilon_n)$ vérifiant

$$\overline{B}(x_n, \epsilon_n) \subset B(x_{n-1}, \epsilon_{n-1}) \cap U_n, \quad \text{avec} \quad \epsilon_n < 2^{-n}$$

La suite (x_n), formée par les centres de ces boules est une suite de Cauchy, puisque si $n \geq m$, alors $\overline{B}(x_n, \epsilon_n) \subset \overline{B}(x_m, \epsilon_m)$ et par suite la distance de x_n à x_m est inférieure ou égale à $\epsilon_m < 2^{-m}$. Donc cette suite converge vers un élément a de E (car E est complet). On vérifie rapidement que a se trouve dans chaque boule fermée $\overline{B}(x_n, \epsilon_n)$, comme celle-ci est incluse dans $B(x, \epsilon) \cap U_n$, on obtient

$$a \in \bigcap_n \overline{B}(x_n, \epsilon_n) \subset B(x, \epsilon) \bigcap_n U_n = B(x, \epsilon) \cap U$$

ce qui achève la démonstration. □

Le théorème de Baire est un des résultats fondamentaux de l'analyse. Il est en effet à la base des théorèmes de Banach et de Banach-Steinhaus.

3.3 Théorème de Banch-Steinhaus. *Soient E et F deux espaces de Banach et (A_n) une suite d'applications linéaires continues de E dans F. Alors les trois propriétés suivantes sont équivalentes*

(a) *Pour tout $x \in E$, la suite $(A_n x)$ a une limite dans F.*

(b) *Il existe une partie totale $X \subset E$ telle que la suite $(A_n x)$ ait une limite dans F pour tout $x \in X$, et il existe une constante C telle que pour tout $n \geq 1$, on ait $\|A_n\| \leq C$.*

(c) *Il existe une application linéaire continue A de E dans F telle que*

$$\lim_{n \to +\infty} A_n x = Ax, \quad \forall x \in E$$

Démonstration. — On va d'abord donner la preuve de la partie facile (et utile) du théorème : $(b) \Longrightarrow (a)$.
Soit E' l'espace vectoriel des combinaisons linéaires finies d'éléments de X. Alors, on définit $A : E' \to F$ par

$$Ax = \lim_{n \to +\infty} A_n x, \quad \forall x \in E'$$

Par passage à la limite dans les inégalités, On a

$$\|Ax\| \leq C\|x\|, \quad \forall x \in E'$$

Le prolongement de A à E, que nous noterons encore par A, s'obtient grâce au théorème 1 de la section 2 et l'inégalité précédente est encore vraie pour tout $x \in E$. Il reste à montrer que

$$\lim_{n \to +\infty} A_n x = Ax, \quad \forall x \in E$$

A cet effet, pour tout x dans E et tout $\epsilon > 0$, on choisit $\zeta \in E'$ tel que $\|x - \zeta\| \leq \epsilon C^{-1}$, il vient

$$\begin{aligned} \|Ax - A_n x\| &\leq \|Ax - A\zeta\| + \|A\zeta - A_n\zeta\| + \|A_n\zeta - A_n x\| \\ &\leq 2\epsilon + \|A\zeta - A_n\zeta\| \end{aligned}$$

Il suffit alors de remarquer qu'il existe n_0 tel que $\|A_n\zeta - A_n\zeta\| \leq \epsilon$ dès que n dépasse n_0.

Venons-en maintenant à l'implication $(a) \Longrightarrow (b)$. Elle est connue sous le nom de théorème de la borne uniforme. En effet, on va montrer que si pour tout $x \in E$ la suite $(A_n x)$ est bornée (ce qui est plus faible que (a)) :

$$\forall x \in E, \text{ il existe } C(x) > 0 \text{ telle que } \|A_n x\| \leq C(x), \ \forall n$$

alors il existe une constante C telle que $\|A_n x\| \leq C\|x\|$, pour tout $x \in E$ et tout n, c'est-à-dire que (A_n) est uniformément bornée par la constante C. Posons $F_k = \{ x \in E;\ \|A_n x\| \leq k, \forall n \}$. Les ensembles F_k sont des fermés dont la réunion est égale à E. Le théorème de Baire assure l'existence d'un F_{k_0} d'intérieur non vide et par suite il existe une boule fermée $\overline{B}(x_0, \epsilon)$ contenue dans F_{k_0}. En particulier

$$\forall n,\ \|A_n x_0\| \leq k_0 \text{ et pour tout } |x| \leq \epsilon,\ \|A_n(x_0 + x)\| \leq k_0$$

On en déduit que, pour tout $x \in \overline{B}(o, \epsilon)$,

$$\|A_n x\| \leq 2k_0, \ \forall n$$

Il suffit alors de prendre $C = 2k_0/\epsilon$. Le reste de la preuve est facile. □

Remarque. — Notons que c'est l'implication $(b) \Longrightarrow (a)$ qui est la plus utile, cela signifie qu'on doit, pour prouver (a), passer obligatoirement par (b). Notons aussi que la preuve montre que l'espace vectoriel F n'a pas besoin d'être complet.

3.4 Théorème de l'application ouverte. *Soient E et F deux espaces de Banach. Si une application linéaire continue A de E dans F est surjective, alors pour tout ouvert U de E, $A(U)$ est un ouvert de F.*

Démonstration. — Désignons par $B_E(o,n)$ la boule dans E, de centre l'origine o et de rayon n. Par hypothèse, on a

$$F = A(E) = A\Big(\bigcup_n B_E(o,n)\Big) = \bigcup_n \overline{A(B_E(o,n))}$$

En vertu du théorème de Baire, il existe n_0 tel que $\overline{A(B_E(o,n_0))}$ contienne une boule $B_F(y_0,\epsilon)$, comme $\overline{A(B_E(o,n_0))}$ est équilibré, il contient aussi $B_F(-y_0,\epsilon)$. De plus, $\overline{A(B_E(o,n_0))}$ est convexe (car l'image d'un convexe par une application linéaire est convexe et la fermeture d'un convexe l'est aussi), il contient donc l'enveloppe convexe des deux boules et par suite contient $B_F(o,\epsilon)$ qui est contenue dans cette enveloppe convexe. Par homothétie, on a

$$\forall r>0,\ \ B_F(o,r) \subset \overline{A(B_E(o,n_0 r\epsilon^{-1}))}$$

en particulier,

$$B_F(o,\epsilon n_0^{-1}) \subset \overline{A(B_E(o,1))}$$

Montrons à présent que $B_F(o,\epsilon/(2n_0))$ est contenue dans $A(B_E(o,1))$. Soit $y \in B_F(o,\epsilon/(2n_0))$, puisqu'on sait déjà que $B_F(o,\epsilon/(2n_0))$ est incluse dans $\overline{A(B_E(o,2^{-1}))}$, il est possible de choisir un point y_1 dans $A(B_E(o,2^{-1}))$ tel que $|y-y_1| < \epsilon/(4n_0)$. Comme $B_F(o,\epsilon/(4n_0))$ est incluse dans $\overline{A(B_E(o,1/4))}$, on peut trouver de même y_2 dans $A(B_E(o,1/4))$ tel que $|y-y_1-y_2| < \epsilon/(8n_0)$. En continuant ce procédé, on construit une suite (y_n) telle que $y_n \in B_F(o,2^{-n})$ et $|y-y_1-y_2-\cdots-y_n| < \epsilon/(2^{n+1}n_0)$. Cela implique que $y=\sum_1^\infty y_n$ et puisque $y_n = Ax_n$, où $x_n \in A(B_E(o,2^{-n}))$, la série de terme général x_n converge, et sa somme x vérifie $\|x\| < 1$, de plus la continuité de A implique que $Ax=y$. Ainsi, on a montré que $B_F(o,\epsilon/(2n_0)) \subset A(B_E(o,1))$.
Par homothétie, on aura aussi $B_F(o,r) \subset A(B_E(o,r\epsilon/(2n_0)))$ et ce pour tout $r>0$. On en déduit que $A(B_E(x_0,r))$ contient la boule $B_F(Ax_0,r\epsilon/(2n_0))$. Ce qui achève la preuve. □

Ce théorème est souvent utilisé dans les situations suivantes :

3.5 Corollaire. *Soit A une application linéaire et bijective d'un espace de Banach E sur un espace de Banach F. Si A est continue, alors A^{-1}, l'application réciproque de A, est elle aussi continue.*

3.6 COROLLAIRE. *Soit E un espace de Banach relativement à la norme $\| \|_1$. Si $\| \|_2$ est une autre norme relativement à laquelle E est aussi un espace de Banach et s'il existe une constante c telle que $\|x\|_2 \leq c\|x\|_1$, pour tout $x \in E$, alors les deux normes sont équivalentes, c'est-à-dire qu'il existe une constante c' telle que $\|x\|_1 \leq c'\|x\|_2$ pour tout $x \in E$.*

Démonstration. — Il suffit d'appliquer le corollaire 5 à l'application identité de $(E, \| \|_1)$ sur $(E, \| \|_2)$. □

3.7 REMARQUES.

• Un raisonnement analogue est valable pour deux familles de semi-normes faisant de E un espace métrique complet : si l'une des deux familles de semi-normes majore l'autre, alors ces deux familles définissent une même topologie. Cette circonstance se rencontre souvent en théorie des distributions.

• Attention, un même espace vectoriel peut être un espace de Banach pour deux normes $\| \|_1$ et $\| \|_2$ sans que ces normes soient équivalentes.

Une autre application intéressante du théorème de l'application ouverte est donnée par le théorème du graphe fermé. Soient E et F deux espaces de Banach. L'ensemble produit $E \times F$ est un espace vectoriel si l'addition et la multiplication par un scalaire sont définies par

$$\alpha(x_1, y_1) + \beta(x_2, y_2) = (\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2)$$

On vérifie que l'application : $(x, y) \to \|(x, y)\| = \|x\|_E + \|y\|_F$ est une norme sur $E \times F$ qui en fait un espace de Banach. Soit $A : E \to F$ une application linéaire. Le graphe de A est le sous-ensemble G_A de $E \times F$ donné par

$$G_A = \{ (x, Ax);\ x \in E \}$$

Puisque A est linéaire, G_A est un sous-espace vectoriel de $E \times F$ et

$$\|(x, Ax)\| = \|x\|_E + \|Ax\|_F$$

Il est facile de voir que si A est continue, G_A est fermé donc complet, le théorème du graphe fermé montre que la réciproque est vraie, plus précisément

3.7 THÉORÈME DU GRAPHE FERMÉ. *Soient E et F deux espaces de Banach, $A : E \to F$ une application linéaire et G_A le graphe de A. Les assertions suivantes sont équivalentes*

(a) *L'application A est continue,*
(b) *Le graphe de A est fermé dans $E \times F$,*
(c) *Si (x_n) est une suite dans E telle que les limites*

$$x = \lim_{n\to\infty} x_n \quad \text{et} \quad y = \lim_{n\to\infty} Ax_n$$

existent, alors $y = Ax$.

Démonstration. — Il est facile de vérifier que (b) et (c) sont équivalentes et que (a) implique (b). Il reste à montrer que (b) implique (a). Pour cela, soient $\pi_1 : G_A \to E$ et $\pi_2 : E \times F \to F$ les projections canoniques, définies par

$$\pi_1(x, Ax) = x, \quad \pi_2(x, y) = y$$

L'application π_1 est linéaire continue et surjective (en fait bijective) de G_A (qui est un espace de Banach d'après (b)) sur E. Le théorème de l'application ouverte montre alors que π_1^{-1} est continue. Or, $A = \pi_2 \circ \pi_1^{-1}$ et π_2 est continue, donc A est continue. □

4. Le théorème d'Arzelà-Ascoli

On sait que la convergence uniforme d'une suite (f_n) de fonctions vers une fonction f implique la convergence simple de (f_n) vers f, alors que la convergence simple de (f_n) vers f n'implique pas nécessairement la convergence uniforme. Par exemple les fonctions $f_n : [0, 1] \to \mathbb{R}$, définies par $f_n(x) = x^n$ converge simplement vers la fonction f définie par

$$f(x) = \begin{cases} 0, & \text{si } 0 \le x < 1 \\ 1, & \text{si } x \le 0 \end{cases}$$

mais la convergence n'est pas uniforme (puisque f n'est pas continue!).

Soit E un espace métrique compact et F un espace métrique complet. On désigne par $C(E, F)$ l'espace des fonctions continues de E dans F et on désigne par d_∞ la distance de la convergence uniforme

$$d_\infty(f, g) = \sup\{\, d(f(x), g(x)); x \in E\,\}$$

c'est un espace complet. Le théorème d'Arzelà-Ascoli caractérise les sous-ensembles compacts de $C(E, F)$ et permet donc de dire quand est ce

qu'on peut en extraire des sous-suites uniformément convergentes. Avant de l'énoncer, on rappelle d'abord la notion d'équicontinuité d'ensemble de fonctions continues.

4.1 DÉFINITION. *Un sous-ensemble $\mathcal{F}$ de $C(E,F)$ est équicontinu en x_0 si, pour tout $\epsilon > 0$, il existe $\eta := \eta(\epsilon, x_0)$ tel que*

$$\forall f \in \mathcal{F},\ d(x, x_0) \leq \eta \Longrightarrow d(f(x), f(x_0)) \leq \epsilon$$

le sous-ensemble $\mathcal{F}$ est dit équicontinu s'il est équicontinu en tout point de E. Il est uniformément équicontinu si η ne dépend que de ϵ.

4.2 THÉORÈME D'ARZELÀ-ASCOLI. *Soient E un espace métrique compact, F un espace métrique complet et $\mathcal{F}$ un sous-ensemble de $C(E,F)$. Pour que $\mathcal{F}$ soit un sous-ensemble compact de $C(E,F)$, il faut et il suffit que $\mathcal{F}$ soit fermé, équicontinu et que, pour tout x dans E, les ensembles $\mathcal{F}(x) = \{ f(x); f \in \mathcal{F} \}$ soient relativement compacts dans F.*

Démonstration. — 1) Supposons $\mathcal{F}$ compact. Il est alors fermé. D'autre part, si $x \in E$, l'application qui à $f \in \mathcal{F}$ associe $f(x) \in F$ est continue puisque

$$d(f(x), g(x)) \leq d_\infty(f,g), \quad f, g \in C(E,F)$$

Comme $\mathcal{F}(x)$ est l'image du compact $\mathcal{F}$ par cette application, il est lui-même un sous-ensemble compact de F. Enfin, puisque $\mathcal{F}$ est compact, pour tout $\epsilon > 0$, il existe un nombre fini de fonctions $f_1, f_2, \ldots, f_N$ dans $\mathcal{F}$ telles que tout f de $\mathcal{F}$ se trouve à une distance inférieure ou égale à $\epsilon/3$ de l'une de ces fonctions. Puisque les f_j sont continues, pour tout $x_0 \in E$, il existe $\eta := \eta(\epsilon, x_0)$ tel que

$$\max_{1 \leq j \leq n} d(f_j(y), f_j(x_0)) \leq \epsilon/3, \quad \text{dès que} \quad d(x_0, y) \leq \eta$$

Soit $f \in \mathcal{F}$, il existe j compris entre 1 et N, tel que $d_\infty(f, f_j) \leq \epsilon/3$ et on en déduit que, si $d(x_0, y) \leq \eta$, alors

$$\begin{aligned} d(f(x_0), f(y)) &\leq d(f(x_0), f_j(x_0)) + d(f_j(x_0), f_j(y)) + d(f_j(y), f(y)) \\ &\leq 2d_\infty(f, f_j) + d(f_j(x), f_j(y)) \leq \epsilon \end{aligned}$$

Cela montre que $\mathcal{F}$ est équicontinu.

2) Inversement, supposons $\mathcal{F}$ fermé et équicontinu et que, pour tout x dans E, $\mathcal{F}(x)$ est relativement compact. Soit $\epsilon > 0$ fixé, l'équicontinuité montre qu'à tout $x \in E$ on peut associer un nombre $\eta := \eta(\epsilon, x)$ tel que

$$d(x, y) \leq \eta \Longrightarrow d(f(x), f(y)) \leq \epsilon/3, \quad \forall f \in \mathcal{F}$$

Puisque E est compact, on peut le recouvrir par un nombre fini de boules ouvertes $B(x_j, \eta)$ $(1 \leq j \leq p)$. Comme $\mathcal{F}(x)$ est relativement compact pour tout x, on en déduit que l'ensemble $\{ (f(x_1), f(x_2), \ldots, f(x_p));\ f \in \mathcal{F} \}$ est relativement compact dans F^p. Il peut donc être recouvert par un nombre fini de boules de rayon $\epsilon/3$, c'est-à-dire qu'il existe $f_1, f_2, \ldots, f_k$ dans $\mathcal{F}$, telles que

$$\mathcal{F} \subset \bigcup_{i=1}^{k} B(f_i, \epsilon/3)$$

Fixons $f \in \mathcal{F}$, et soit i tel que f soit dans $B(f_i, \epsilon/3)$. Si $x \in E$, alors il existe j tel que $x \in B(x_j, \eta)$, et par suite

$$d(f(x), f_i(x)) \leq d(f(x), f(x_j)) + d(f(x_j), f_i(x_j)) + d(f_i(x_j), f_i(x)) \leq \epsilon$$

ce qui implique que $d_\infty(f, f_i) \leq \epsilon$. Cela, joint au fait que $\mathcal{F}$ est fermé donc complet, implique que $\mathcal{F}$ est compact. □

Il existe d'autres situations où la vonvergence simple d'une suite de fonctions continues, vers une fonction continue, implique la convergence uniforme. En voici une, imaginée par Dini

4.3 PROPOSITION. *Supposons E compact et F complet et soit (f_n) une suite de fonctions continues convergeant simplement vers une fonction continue g. Si la suite (f_n) satisfait la propriété suivante : il existe une constante $c \geq 1$ telle que*

$$\forall m \geq n \geq 1, \quad d(g(x), f_m(x)) \leq c.d(g(x), f_n(x))$$

alors, (f_n) converge uniformément vers g.

Démonstration. — Soit $\epsilon > 0$ fixé. La convergence simple implique que pour tout x dans E, il existe un entier $N(x)$ tel que $d(f(x), f_n(x)) \leq \epsilon/(3c)$

dès que n dépasse $N(x)$. Puisque f et $f_{N(x)}$ sont continues en x, il existe $\eta(x)$ tel que pour tout $y \in E$, vérifiant $d(x,y) \leq \eta(x)$, on ait

$$(*) \qquad \max\{\, d(f(x), f(y)),\ d(f_{N(x)}(x), f_{N(x)}(y))\,\} \leq \epsilon/(3c)$$

Comme E est compact, il existe un nombre fini de boules ouvertes $B(x_i, \eta(x_i))$, avec $(1 \leq i \leq n)$, telles que $E \subset \bigcup_{i=1}^{n} B(x_i, \eta(x_i))$. Posons $N_0 = \max\{\, N(x_i),\ 1 \leq i \leq n\,\}$ et soit $n \geq N_0$ quelconque. Pour tout élément x de E, il existe x_i tel que x soit dans $B(x_i, \eta(x_i))$. La propriété que vérifient la suite (f_n) et l'inégalité triangulaire impliquent que

$$\begin{aligned} d(f(x), f_n(x)) &\leq c.d(f(x), f_{N(x_i)}(x)) \leq \\ &\leq c\Big[d(f(x), f(x_i)) + d(f(x_i), f_{N(x_i)}(x_i)) + d(f_{N(x_i)}(x_i), f_{N(x_i)}(x))\Big] \end{aligned}$$

Comme x est dans $B(x_i, \eta(x_i))$, l'inégalité $(*)$ montre que

$$d(f(x), f(x_i)) \leq \epsilon/(3c) \quad \text{et} \quad d(f_{N(x_i)}(x), f_{N(x_i)}(x_i)) \leq \epsilon/(3c)$$

de plus, puisque $n \geq N_0 \geq N(x_i)$, on a aussi $d(f(x_i), f_n(x_i)) \leq \epsilon/(3c)$. De cette inégalité et celles qui la précèdent, on déduit que, pour tout $x \in E$ et pour tout $n \geq N_0$, $d(f(x), f_n(x)) \leq \epsilon$, c'est-à-dire que (f_n) converge uniformément vers f. □

4.4 Théorème de Dini. *Soit E un espace métrique compact. Si (f_n) est une suite monotone de fonctions continues, convergeant simplement vers une fonction continue g, alors (f_n) converge uniformément vers g.*

Démonstration. — C'est une conséquence immédiate de la proposition précédente. En effet, si la suite (f_n) est, par exemple, croissante c'est-à-dire $f_{n+1} \geq f_n$, et si $m \geq n$, alors

$$|f(x) - f_m(x)| \leq |f(x) - f_n(x)|, \quad \forall x \in E$$

La suite (f_n) satisfait donc les hypothèses de la proposition 4.3 □

Bibiliographie

[1] A. Achour, *Calcul Différentiel, Cours et Exercices*, Collection M/ Sciences fondamentales, Centre de Publication Universitaire, 1999.

[2] J. P. Aubin, *Initiation à l'analyse appliquée*, Masson 1994.

[3] F. Bayer–C. Margaria, *Espaces de Hilbert et opérateurs*, Tome 2, Ellipse, 1986.

[4] samuel S. Holland Jr., *Applied analysis by Hilbert space method*, Marcel Dekker, Inc, 1990.

[5] A. Kirillov, A. Gvichiani, *Théorèmes et problèmes d'analyse fonctionnelle*, Traduction française, Editions Mir, 1982.

Cet ouvrage est une introduction progressive à l'analyse hilbertienne et à la théorie spectrale des endomorphismes continus d'espaces de Hilbert. Il s'adresse aux étudiants de deuxième cycle préparant une maîtrise de mathématiques, de mathématiques appliquées ou de physique, aux candidats à un concours de recrutement, CAPES ou Agrégation, ainsi qu'aux élèves-ingénieurs qui pourront donc l'utiliser avec profit.

Le livre est organisé en cinq chapitres et une annexe. Chaque chapitre est divisé en sections et chaque section est suivie d'une liste d'exercices de difficulté variée, permettant de contrôler la bonne assimilation du cours. L'annexe contient les principaux théorèmes d'analyse fonctionnelle, utilisés dans ce cours, ainsi que leurs démonstrations.

Une biographie succinte, puisée dans l'Encyclopedia Universalis et relative aux mathématiciens qui ont donné leur nom aux théorèmes évoqués dans cet ouvrage, permet de situer ces théorèmes dans leur contexte chronologique.

Houcine Chebli est docteur ès-sciences mathématiques. Il est professeur à la Faculté des sciences de Tunis depuis 1975 et a enseigné les mathématiques à tous les niveaux de la maîtrise. Son activité de recherche porte sur la théorie spectrale et l'analyse harmonique.

Couverture, ill. ESCHER visions, Doris Schattschneider, SEUIL.



ISBN : 9973-37-013-9 Prix : 12 dinars